

Office of the Chief Scientist for Human Factors

Aviation Maintenance Human Factors

Program Review
FY04



Dr. William K. Krebs, Aviation Maintenance Human Factors Program Manager

**Federal Aviation Administration
ATO-P R&D HF (Room 907A)
800 Independence Avenue, S.W.
Washington, D.C. 20591
phone (202) 267-8758
e-mail: william.krebs@faa.gov
<http://www.hf.faa.gov/krebs>**

Aviation Maintenance human factors research has the overall goal to identify and optimize the factors that affect human performance in maintenance and inspection. The focus initiates on the technician but extends to the entire engineering and technical organizational and all personnel involved in the endeavor. Research attention to personnel can include selection, qualification, training, motivation, health, professionalism, and the variety of human capabilities and limitations that affect efficient and safe maintenance task performance. The research considers many aspects of the work environment including both the physical and social aspects of the organization. The complexity of technical communication is an example of such research. The diversity of maintenance and inspection activity is unlimited. Thus the research attends to each and every action preformed by individuals, teams, departments, and the collective organization. With a view of people, the environment in which they work, and the actions they perform a final focus is on the resources necessary for efficient and safe work. Research related to resources includes studies on the design of documentation and procedures, selection of tools, equipment, buildings, applications of advanced technologies for maintenance and inspection. The maintenance human factors research combines critical basic scientific understanding of human performance with applied studies conducted in cooperation with industry partners. The results are solid and proven science, psychology, and engineering delivered in plans, procedures, software, and even hardware that can be immediately implemented to affect efficiency and safety. To obtain a detailed description of current aviation maintenance human factors projects, projects completed, accomplishments, and products delivered, please point to <http://www.hf.faa.gov/maintenance.htm>. Dr. Bill Johnson is the Chief Scientist of Aviation Maintenance Human Factors and Dr. William “Kip” Krebs is the research program manager.

The following report lists projects between October 1st, 2003 and September 30th, 2004. These projects address requirements identified by the Federal Aviation Administration Flight Standards office. The intent of this report is to allow Federal Aviation Administration sponsors to determine whether their requirements have been satisfactorily addressed, allow investigators to receive feedback from Federal Aviation Administration sponsors and other interested parties, and to provide feedback to the ATO-P R&D HF aviation maintenance program manager on the quality of the research program. Basically, this document is a means of holding each group (sponsor, investigator, ATO-P R&D HF program manager) accountable to ensure that the program is successful.

In FY04, the aviation maintenance research program distributed \$700,000 contract and grant dollars to multiple organizations. In addition, one project received supplemental support from the Civil Aerospace Medical Institute, Oklahoma City, OK.

Address questions or comments to:

William K. Krebs, Ph.D.

Aviation Maintenance Human Factors

FY04 Funded Projects

<u>Project Title</u>	<u>Page #</u>
General Aviation Maintenance Accidents: An Analysis Using HFACS and Focus Groups	4
Factors Affecting the Success or Failure of Aviation Safety Action Programs in Aviation Maintenance Organizations	9
Computer and Broadband Technology in Aircraft Line Maintenance: A Task Analysis and Questionnaire	14
Fatigue and Aircraft Inspection 1: Model and Simulation	21
Year 2: Development of The General Aviation Inspection Training Systems (Gaits)	27
Experiments on Language Errors in Aviation Maintenance: Asia	35
Empirical Basis for An Occupational Vision Standard	40
An Analysis of the Visual Demands Associated with Aviation Maintenance Inspectors	51
Visual Functioning of Aviation Maintenance Inspectors	56
Year1: Development of a Web-Based Surveillance and Auditing Tool to Reduce Aviation Maintenance Errors (WebSat)	61

GENERAL AVIATION MAINTENANCE ACCIDENTS: AN ANALYSIS USING HFACS AND FOCUS GROUPS

Albert Boquet, Ph.D., Cristy Detwiler, B.A., Carrie Roberts, B.A., Dan Jack, M.S., & Scott Shappell, Ph.D.
FAA/Civil Aerospace Medical Institute

Douglas A. Wiegmann, Ph.D.
University of Illinois at Urbana-Champaign

The Human Factors Analysis and Classification System (HFACS) was used to classify maintenance-related general aviation accidents in the United States from 1990 to 2000 inclusive. The analysis revealed that among the maintainers, skill-based errors were most frequent cause of accidents, followed by violations committed by both professional maintainers and owner-operators. Furthermore, violations committed by owner-operators were twice as likely to be associated with a fatality. In addition, focus groups comprised of professional airframe and powerplant mechanics in both Alaska and Oklahoma, provided valuable information to validate the accident analysis and describe the state of general aviation maintenance today.

INTRODUCTION

Commercial carriers have invested a great deal of financial and corporate resources to address human factors both on the flight deck and within maintenance. However, by comparison, general aviation (GA) has lagged somewhat behind. This is surprising when one considers that as much as 96% of active aviation in the United States involves either general or corporate aviation (Wells, 1996). For instance, Ropp and Lopp in 1998, found both general and corporate aviation lacking in any sort of structured safety management system for maintenance operations, in spite of the fact that maintenance related accidents comprised as much as 21.3% of the accidents occurring in 1997. This number is in stark contrast to the 9.7% of maintenance related accidents from 1987 to 1996 reported by Boeing (1997) for commercial aviation. In light of the fact that the accident rate for GA aircraft is five to seven times that of commercial air carriers, these percentages take on even more significance.

That is not to say that nothing has been done at all to address this concern. Indeed, an earlier study of maintenance-related GA accidents conducted by Goldman, Fiedler, & King, (2002) examined 1,503 National Transportation Safety Board (NTSB) accident reports spanning the years of 1988 to 1997. Their findings revealed that the most common accident cause factors involved installation errors, general maintenance, and maintenance inspection. Furthermore, they demonstrated that installation errors were often associated with severe accidents. In fact, their findings indicate that installation problems, general maintenance, and maintenance inspection accounted for over 50% of the fatalities in their sample. While these findings provide valuable evidence for the role of human error in GA maintenance, the results were limited in that the subject matter experts (SMEs) who evaluated the NTSB reports were actually GA pilots and not active aviation maintenance technicians (AMTs).

Likewise, one cannot study the types of errors associated with AMT performance in a vacuum. One must also bear in mind the environment within which the errors occur. For example, a majority of maintenance inspection is visual. This necessitates adequate lighting in the workplace, be that workplace indoors such as in a standing structure, or outdoors,

where one may assume a fair portion of GA maintenance might occur. Indeed, AMTs are often required to work in less than optimal environments which may include one or some combination of unsafe noise levels, heat, cold, poor lighting and restricted workspace. Thus, one cannot exclude the environmental component associated with aircraft maintenance.

With this in mind, it makes sense to not only try to understand the errors made within the context of GA maintenance, but environmental factors as well. This line of reasoning has led experts and government agencies such as the Federal Aviation Administration (FAA) to examine not only the underlying factors involved in GA accidents, but to specifically target GA accidents in Alaska, a region known for its harsh climate and environmental conditions.

Consequently, the FY04 maintenance human factors effort at the Civil Aerospace Medical Institute (CAMI) had two purposes. One was to investigate human error associated with GA maintenance related accidents. The second purpose was to compare the errors made in Alaska (AK) with the rest of the United States (RoUS). To this end, not only were the maintenance factors associated with GA accidents investigated but focus group interviews of AMTs both in AK and in Oklahoma were conducted in an attempt to define issues faced by GA AMTs both in Alaska and at least one site in the contiguous 48 states.

HFACS

The entire HFACS framework includes a total of 19 causal categories within Reason's (1990) four levels of human failure. While in many ways, all of the causal categories are equally important; particularly germane to any examination of GA accident data are the unsafe acts of aircrew. For that reason, we have elected to restrict this analysis to only those causal categories associated with the unsafe acts of GA aircrew. A complete description of the HFACS causal categories is therefore beyond the scope of this report and can be found elsewhere (Wiegmann & Shappell, 2003).

Unsafe Acts of Operators

In general, the unsafe acts of operators (in the case of aviation, the aircrew) can be loosely classified as either errors

or violations (Reason, 1990). Errors represent the mental or physical activities of individuals that fail to achieve their intended outcome. Not surprising, given the fact that human beings by their very nature make errors, these unsafe acts dominate most accident databases. Violations on the other hand, are much less common and refer to the willful disregard for the rules and regulations that govern the safety of flight.

Within HFACS, the category of errors was expanded to include three basic error types (decision, skill-based, and perceptual errors). In general, decision errors represent conscious decisions/choices made by an individual that are carried out as intended, but prove inadequate for the situation at hand. In contrast, skill-based behavior within the context of aviation is best described as “stick-and-rudder” or other basic flight skills that occur without significant conscious thought. As a result, these skill-based actions are particularly vulnerable to failures of attention and/or memory as well as simple technique failures. Finally, perceptual errors occur when sensory input is degraded or “unusual,” as is often the case when flying at night, in the weather, or in other visually impoverished conditions.

While errors occur when aircrews are behaving within the rules and regulations implemented by an organization, violations represent the willful disregard for the rules and regulations that govern safe flight. As with errors, there are many ways to distinguish between types of violations. However, two distinct forms are commonly referred to, based upon their etiology. The first, routine violations, tend to be habitual by nature and are often tolerated by the governing authority. The second type, exceptional violations, appear as isolated departures from authority not necessarily characteristic of an individual’s behavior nor condoned by management.

METHODS

Data

The National Aviation Safety Data Analysis Center (NASDAC) and NTSB were utilized to identify maintenance related GA accidents. Two methods were used to select the maintenance factor sample. First, a sample of causal factors was selected from the years 1990-2000 based on NTSB personnel codes that identified the involvement of maintenance personnel (i.e., 4107 - Company Maintenance Personnel and 4108 - Other Maintenance Personnel). Second, NTSB “subject” codes were scanned to identify any accidents that involved maintenance causal factors (24100-24124). This latter method was used to ensure that all maintenance factors were captured, including those that were not attributed to a certified AMT or otherwise designated maintainer (e.g., an owner/operator).

Subject Matter Experts

The maintenance causal factors associated with each maintenance related accident were classified into HFACS categories independently by six certified, instructor level airframe and powerplant mechanics (A/P) who served as

mechanic SMEs. The combined years in the aviation industry for the SMEs was 168 years with an average of 28 years. In addition, all were maintenance instructors at a local school. The span of instructor level teaching as aviation mechanics was 3 to 14 years with an average of 8 years.

SME Training

Training in HFACS for the mechanic SMEs was conducted in three phases.

Phase 1: An HFACS training session was conducted by the authors for the purpose of introducing the SMEs to the HFACS framework (Wiegmann and Shappell, 2001) and instructing them on how to use it. From the sample of maintenance related accidents (n=1935), a 10% random sample (n=194) was selected, resulting in 206 maintenance factors to be coded. Together, all six SMEs coded 59 factors from the first 50 accidents and discussed their codes in detail. In three subsequent meetings the remaining factors from the random sample were coded independently by all six SMEs. Initial coder agreement was not computed for this initial phase.

Phase 2: Maintenance factors from the years 1990-1991 were then randomly assigned to pairs of SMEs for coding. Using pairs of coders allowed for analysis of initial coder agreement. The SMEs coded their assigned factors independently. Codes were entered, discrepancy reports were generated, and initial coder agreement was computed. The SMEs agreed approximately 51% of the time during this second phase. Recall however, that there were 19 possible HFACS categories that the SMEs could place the causal factor in, which makes the percentage agreement appear more reasonable. Still, the inter-rater reliability is low when compared with the over 85% level of agreement seen with pilot SMEs coding aircrew errors associated with GA accidents (Wiegmann & Shappell, 2003). All the same, any factor for which the two SME coders had a discrepancy was discussed and resolved by all six SMEs as a group. These group discussions were also used to develop the exemplars within the causal categories associated with the HFACS framework.

Phase 3: This phase was initiated because of lower than anticipated initial coder agreement in Phase 2. Maintenance factors from the years 1999-2000 were coded and resolved using the same methodology as was used in Phase 2. Initial coder agreement increased to 59% for those years. However, it was determined that this percentage was still not high enough to justify the resolution of discrepancies with only two coders as was originally planned. It was therefore decided that the remaining data would be coded and resolved as they had been in Phases 2 and 3 of training, with two independent coders for each factor, and group resolution of discrepancies.

HFACS Coding

After completion of Phase 3 of training, the SMEs coded maintenance factors for the years 1992-1993. The necessity of meeting with all six SMEs to resolve discrepancies was time-

consuming and slowed the coding process considerably. It was decided, in the interest of time and completeness, that the remaining years of data (1994-1998) would be coded in two separate groups. This allowed a cross-section of data from all years to be analyzed before all of the coding was complete.

Upon completion of the first group for years 1994-1998, the SMEs raised concerns about the reliability and validity of the data obtained from Phases 1 and 2 of training. Therefore, the data coded in these phases were eliminated from the analysis, and were re-coded by the SMEs. Maintenance factors from the years 1990-1991 were also separated into two groups to be coded again. To date, 1263 maintenance causal factors associated with 1133 accidents have been coded (*note: the aircrew and other human causal factors have been coded and reported in previous reports – for a summary see the HFACS FY04 Annual Report*).

GA Maintenance Focus Group

In order to better understand the issues facing maintenance providers in Alaska today, and to validate the HFACS analysis, a series of focus groups were conducted at selected maintenance sites throughout Alaska. These focus groups were composed of personnel at maintenance facilities located in Anchorage, Nome, Fairbanks, Juneau, and Barrow, Alaska. The results of these interviews were then compared with focus group interviews made up of the SMEs in Oklahoma City, OK.

RESULTS

HFACS Analysis

Similar to other areas of aviation, skill-based errors (SBEs) were associated with the largest percentage of maintenance related accidents (Wiegmann & Shappell, 2003; Figure 1). These types of errors were followed by violations committed by AMTs (VMAINT) at 23.9%, violations by owner/operators (VOO) at 12.1% and decision errors (DE) at 8.2%. Of note, no perceptual errors were reported by the SMEs for maintenance related data.

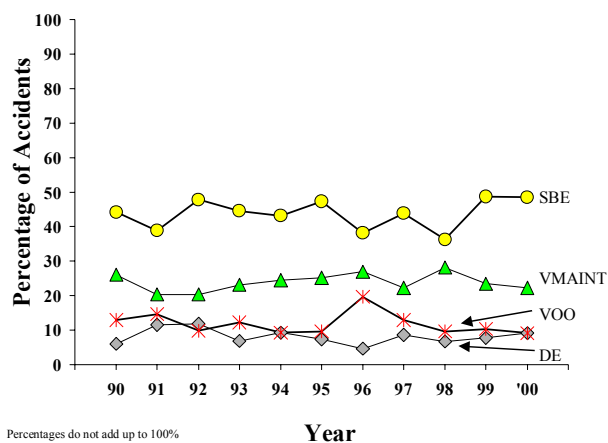


Figure 1. Overall accident rate by year and unsafe act.

Fine-Grained Analysis

In order to gain a better understanding of the specific types of errors committed, a fine-grained analysis was conducted for each of the unsafe acts reported above. Those errors, which comprised at least 5% of the unsafe acts within each HFACS error category, were reported. A brief summary of those results follows:

Decision Errors. The most common decision error was the failure to comply with a service bulletin or letter. This comprised 35.2% of the decision errors in the sample. These decision errors were followed by maintenance overhaul (11.2%), and replacement of parts (8.0%).

Skill-Based Errors. The fine-grained analysis revealed that the most common skill-based error was installation, which accounted for 29.3%, followed by inspection errors accounting for 16.7%.

Aviation Maintenance Technician Violations. Violations attributed to AMTs were similar to skill-based errors in that the most common violation involved installation (16.7%), while the failure to follow procedures and directives were the second highest violation committed by an AMT at 12.6%.

Owner/Operator Violations. Violations committed by owner-operators performing their own maintenance were somewhat different from those committed by AMTs. The most common violation in this case was the failure to obtain an annual inspection (18.2%). Following that, aircraft service and maintenance represented the next highest percentage of violations seen with owner/operators (10.6% each). Improper installation resulted in 10.9% of the violations, and unauthorized design change, modifications, and non-compliance with airmen's directives each accounted for 5.2% of violations observed in this causal category.

Comparison between Alaska and the Rest of the U.S.

Because of the disparity in total events between AK and the RoUS, the comparison between the two will reflect aggregate numbers collapsed across the 10-year period rather than an annual comparison. This was done to account for the relatively small cell sizes found in the AK data.

The percentage of skill-based errors associated with maintenance related accidents for AK and the RoUS were essentially the same (AK=43.4%; RoUS=46.7%). Similar patterns were noted for decision errors with 8.1% of the maintenance-related accidents in AK associated with decision errors versus 11.2% in the RoUS. Likewise, violations for both AMTs and owner-operators revealed almost identical patterns whether they occurred in AK or the RoUS (AK = 23.9%, RoUS = 22.2% and AK = 12.1%, RoUS = 13.3%).

Fatal Events Related to Maintenance Unsafe Acts

In an effort to quantify a worst-case scenario of maintenance-related accidents, the unsafe acts were examined with respect to the degree that they factored into a fatal event.

The percentage of fatal and non-fatal maintenance related accidents associated with each of the unsafe acts is presented

in Figure 2. What is evident is that skill-based errors are least likely to be associated with fatal accidents while violations attributed to owner/operators were most often associated with fatal accidents by an almost 3 to 1 margin. Indeed, nearly 1/3 of the accidents attributed in part to a maintenance violation committed by an owner/operator were associated with fatalities. Decidedly, fewer fatalities were attributed to violations committed by AMTs, although even they were twice as likely to result in fatalities when compared with skill-based errors.

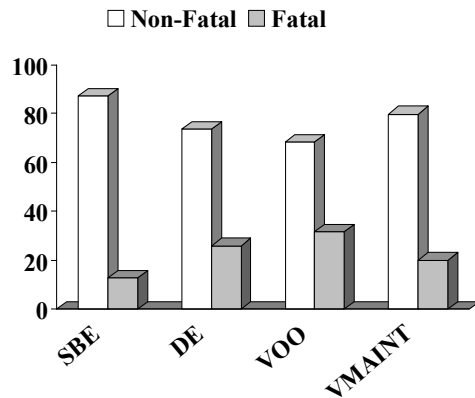


Figure 2. Percentage of maintenance related unsafe acts associated with fatal and non-fatal accidents.

Maintenance Focus Group Analysis

In an effort to understand the issues facing AMTs in today's GA environment, a series of focus group surveys were carried out both in AK and in OK. Although far from complete, this initial effort was initiated to get a better understanding of those areas of GA maintenance that need to be addressed both from a regulatory, as well as from a maintenance/system safety, standpoint. Further interviews are planned for other regions of the U.S. in FY05.

That being said, the data obtained from Alaska and Oklahoma were revealing and will be briefly summarized here.

Alaska. A number of problems were mentioned by the Alaskan focus groups, ranging from training programs to oversight (or lack thereof) by regulatory agencies. One area of concern mentioned by our focus group members was licensing. Separate licensing for large aircraft, GA, and rotorcraft, in addition to doing away with endorsements was one possible remedy mentioned. Presumably, this would open the door for advanced training and recognize maintainers for the professionals that they are, not just technicians.

Also obtained from the focus groups was the apparent lack of qualified maintenance personnel in Alaska. A number of reasons were cited for this with the distinct lack of training facilities topping the list. Poor remuneration for GA maintenance personnel also makes retention difficult. Also of concern was the fact that training beyond certification is hard to come by in Alaska, not to mention expensive. Lack of training in basic mechanics in technical programs was also

cited as a problem. Finally, the focus groups suggested that the pressure to graduate students from programs results in teaching to certification exams, rather than focusing on core subject matter.

Oklahoma. The focus group established in the Oklahoma City area echoed many of the same sentiments of the Alaska focus groups. For instance, the group was unanimous in their assertion that there were not enough qualified GA mechanics to meet industry demands. Furthermore, they also cited training as a major shortcoming in the industry. Specifically, a lack of training facilities and lack of ongoing training and certification opportunities in the GA sector were a major concern.

Oversight by the FAA was also voiced as a concern by the Oklahoma focus group. In addition, follow-up on manuals once they are submitted, surveillance of pilots performing their own maintenance, and oversight of maintenance performed on weekends and after hours were all cited as issues. Finally, they were concerned that pay rates for GA mechanics were too low, which might make it difficult to keep people in the field.

DISCUSSION

A number of errors were classified using the HFACS framework including not only AMTs, but also owner-operators performing their own maintenance. Perhaps most notable were violations. For instance, violations committed by AMTs represent an inordinately high percentage of the unsafe acts when compared to violations committed by flight crews (Shappell & Wiegmann, 2003). Moreover, owner-operator violations proved to be an even greater problem in GA maintenance. This observation is supported by the fact that accidents, which were associated with owner-operator violations, were three times more likely to involve a fatality than accidents involving skill-based errors. The data for violations committed by AMTs did not prove to be much better, revealing a two-fold increase in the likelihood of a fatality.

Even more important is determining why the higher percentages of violations occurred in the first place. For the owner-operators, the two most common violations were the failure to obtain an annual inspection and aircraft service/maintenance. Thus, for the owner, it may be the expense of obtaining an inspection and servicing the aircraft, which may cause the owner to delay these services. This makes sense when one considers scheduled maintenance for the family automobile. It's quite likely that manufacturer scheduled maintenance is either not followed to the letter or ignored entirely by those who simply can't afford it. However, as an individual's income improves later in life, so to does the frequency of scheduled maintenance on the family car.

On the other hand, violations attributable to AMTs tended to reflect the business of actually maintaining the aircraft. Specifically, the two most common violations for AMTs were installation and failure to follow procedures and directives.

The fix for this may involve finding a different way to perform certain tasks, which differ from protocols laid out in service manuals or bulletins. The “I know best” mentality may work well in some instances, but has the potential for catastrophe as demonstrated by the data reported here.

Similar to other areas of aviation, the most common unsafe act seen in the maintenance data was skill-based errors. This remains a consistent finding in the analysis of accidents using the HFACS framework, and more than likely is explained by the fact that even in complex environments, the bulk of the behaviors performed by operators tend to be low processing, highly automatized behaviors. However, these findings differ in that there were decidedly fewer skill-based errors noted in the maintenance data than is typically seen in other industrial settings such as aviation and mining. In fact, when one surveys the literature regarding flight crews, the percentage for skill-based errors is approximately double that noted here (Shappell & Wiegmann, 2003). Exactly why this would be the case is hard to say. However, it may be inherent to the job of the AMT where one would expect to find less routine behavior than in the cockpit.

For skill-based errors, both focus groups mentioned a number of interventions that may prove beneficial when addressing the errors and violations observed in our data. For instance, something as simple as ensuring that AMTs have the proper tools to perform tasks would likely enhance technical applications. Training in shift scheduling and the importance of sleep requirements might also help to combat fatigue and related mistakes. Finally, proper lighting and organization of the workspace has been shown to be effective in improving proficiency.

Dealing with violations may prove to be the most difficult of the unsafe acts to address. First, this is not an error *per se*, but willful behavior that is committed by the person charged with insuring that the aircraft is safe to fly. Thus, the same interventions that may prove useful in mitigating human error, don't really apply here. This is perhaps where regulatory agencies may play the most important role. Fair and consistent punitive actions taken against those individuals who violate the rules have been shown to be successful amongst pilots in the U.S. Navy and Marine Corps (Shappell, Wiegmann, Fraser, Gregory, Kinsey, and Squier, 1999). Although policing maintenance operations may prove difficult for any one entity to do, (e.g., the FAA); consistent enforcement may help to send the message that the regulatory agency takes violations as a serious affront to aviation safety.

However, one must also question the safety culture in which these violations occur. Just as GA pilots must be made part of a culture of safe flight, so must those individuals who choose to maintain their aircraft. This culture or attitude of safety begins with the first day of training and should be stressed throughout one's career. In effect, safety begins with the AMT, long before any pilot leaves the ground. So shouldn't the same emphasis be placed on ensuring safety in maintenance operations as is seen in the cockpit?

When comparing the responses of the focus groups, there were far more similarities than differences. In fact, for both groups, the chief complaints were lack of pay, which causes a shortage of personnel in the field. Both groups also cited poor training programs, both for certification and for supplementary training following licensure. Until these issues are addressed, it will be difficult to address any other problems from the AMT side of the equation. Finally, while there was consensus between the focus groups, it should be noted that there were only two regions surveyed. Future work will involve regional focus groups from the rest of the United States.

These data suggest that rather than using a blanket, one-size fits all approach to rectifying these problems, targeted interventions should be employed that will be most effective in reducing the specific types of errors seen here. For example, decision errors, especially those that are knowledge-based, would benefit most from additional on-going training. Furthermore, stressing the importance of following service bulletins and manufacturers maintenance recommendations may influence decision making in the right direction. In fact, by making service bulletins a requirement, would remove the decision-making from the maintainer altogether.

Nevertheless, while interventions and recommendations can be talked about and instituted by employers and regulatory agencies, ultimately, the person holding the wrench has to want to be safe. Only then will they invest themselves in their work and in the safety of the planes that we fly.

REFERENCES

- Boeing. (1997). *Statistical summary of commercial jet airplane accidents*. Boeing Commercial Airplane Group. Seattle, WA, 21.
- Goldman, S., Fiedler, E., & King, R. (2002). *General aviation maintenance-related accidents: A review of ten years of NTSB Data*. Washington, DC: FAA, Office of Aerospace Medicine. Report No: DOT/FAA/AM-02/23.
- Ropp, T. & Lopp, D. (1998). *Human factors in general aviation*. Midway Airport Tenant Association. Midway Airport, Chicago, IL.
- Shappell, S., Wiegmann, D., Fraser, J., Gregory, G., Kinsey, P., & Squier, H. (1999). Beyond mishap rates: A human factors analysis of U.S. Navy/Marine Corps TACAIR and rotary wing mishaps using HFACS. *Aviation, Space, and Environmental Medicine*, 70, 416-417.
- Shappell, S. & Wiegmann, D. (2003). *A human error analysis of general aviation controlled flight into terrain accidents occurring between 1990-1998*. Washington, DC: FAA, Office of Aerospace Medicine. Report No: DOT/FAA/AM-03/4.
- Wells, A. (1996). *Airport planning management*. McGraw, NY. 41.
- Wiegmann, D. & Shappell, S. (2003). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Burlington, VT: Ashgate Publishing Ltd.

FACTORS AFFECTING THE SUCCESS OR FAILURE OF AVIATION SAFETY ACTION PROGRAMS IN AVIATION MAINTENANCE ORGANIZATIONS

Manoj S. Patankar, Ph.D. & David Driscoll
Saint Louis University
St. Louis, Missouri

The objective of the Federal Aviation Administration's (FAA) Aviation Safety Action Program (ASAP) is to encourage air carrier and repair station employees to voluntarily report errors that may be critical to identifying potential precursors to accidents. Under an ASAP, safety issues are resolved through proactive action rather than through punishment or discipline. The goal of this study was to identify factors that may lead to the success or failure of an ASAP. The Maintenance ASAP Questionnaire (MAQ) was developed and distributed to a randomly selected sample of 83,000 certificated aircraft mechanics. The results of this survey indicate that there is an overwhelming belief among the respondents that the ASAP programs can truly improve safety. The hurdles in building a successful ASAP program are rooted in two key areas: (a) limited interpersonal trust among mechanics, managers, and the FAA inspectors and (b) lack of awareness about the ASAP programs as well as its potential benefits. In addition to higher levels of trust and awareness among the organizations with successful ASAP programs, it was also clear that these organizations had a more collaborative labor-management relationship.

INTRODUCTION

In 1996, Aviation Safety Action Programs (ASAPs) were introduced in the flight domain with the hope of encouraging pilots to disclose their errors and, more importantly, the factors contributing to their errors. With this knowledge, systemic solutions could then be implemented to preclude recurrence (Harper & Helmreich, 2003). In the absence of specific disclosure by pilots, vital information is not available to the air carrier or the Federal Aviation Administration (FAA) and the solutions are not likely to be systemic. In order to encourage pilots to participate in such a program, the FAA developed specific guidance (AC 120-66) for all the parties involved: FAA field inspectors, pilots unions, and air carrier management. As delineated in this guidance material, the FAA is genuinely interested in obtaining safety-related information through this non-punitive program. Generally, air carriers with ASAP programs are very satisfied with their programs and they believe that the program has identified systemic discrepancies that would not have been otherwise discovered.

In an effort to expand the scope of the ASAP programs, the FAA added guidance materials for the maintenance community (AC 120-66A and -66B). Prior to the start of this study, there were twenty-eight air carriers with flight ASAP programs and only six organizations with maintenance ASAP programs. Since the beginning of this study, the number of flight ASAP programs has risen to forty-one and the number of maintenance ASAP programs has risen to ten. Although both programs have increased during the past year, the ratio of flight-to-maintenance programs remains steady at about four-to-one.

In terms of the events reported to the respective Event Review Committees (ERCs), the ratio seems to be about ten-to-one: flight ASAPs receive about ten times as many reports as maintenance ASAPs. Nonetheless, due to the "networked" environment in maintenance versus the "linear" environment in flight (Patanekar & Driscoll, 2004), the resources required to investigate and manage the two programs are about the same.

For the purpose of this study, a "successful" ASAP program is defined as the one that has matured to such a level that there is a regular flow of ASAP reports, there are personnel dedicated to maintaining, analyzing, and implementing of these reports, and there is a mechanism established to provide feedback regarding the overall effects or impacts of the ASAP program. Some "highly successful" programs are able to leverage the benefits of similar agreements in their flight, dispatch, and/or cabin crews. An unsuccessful or "failed" ASAP program is defined as a condition wherein there is no signed MOU between the company, labor union, and the FAA regarding an ASAP program—basically, the program does not exist.

The FAA, the maintenance organizations, and the labor unions want to minimize maintenance errors and improve safety. With this ultimate goal in mind, the present study identifies some of the key factors that are likely to lead to a successful ASAP program in aviation maintenance as well as factors that may be preventing them from getting started.

LITERATURE REVIEW

Early interest in proactive non-punitive measures is evident in the Maintenance Resource Management Roundtables conducted at US Airways (Taylor &

Christensen, 1998). An MRM Roundtable, as it was called, consisted of a representative from the company, a representative from the International Association of Machinists and Aerospace Workers, the FAA Principal Maintenance/Avionics Inspector, and the mechanic(s) who committed the error. The tripartite team (FAA, company, and labor union) endeavored to steer clear of the prevalent blame culture (cf. Marx and Graeber, 1994) and sought a better understanding of the causal factors leading to the error. By adopting this approach, the team was successful in winning the labor force's trust and truly implementing comprehensive and systemic solutions. In response to such a program, several key issues were resolved without resulting in an FAA enforcement action against the mechanic or the company. Unfortunately, the roundtable system was practiced at only one company and was difficult to duplicate at other companies because other people (including FAA inspectors and company managers) were not as amenable to such a system. (Taylor & Christensen, 1998).

Mechanics who did not have access to a roundtable discussion, may have had at least two other options: they could either submit a report to NASA's Aviation Safety Reporting System (ASRS) or use the guidance provided in Advisory Circular 00-58 (cf. FAA, 1998) to file a voluntary self-disclosure report. The ASRS report may provide limited protection to the individual reporter, but the reporter's complaint cannot be acted upon by the company management or the FAA because the individual reports are de-identified; however, NASA will provide statistical information to the FAA if a significant number of reports identify the same problem. A self-disclosure report filed in accordance with AC 00-58, on the other hand, will provide additional legal protection and bring the reporter's concern directly to the company management and the FAA. This advisory circular is designed for a generic (not limited to maintenance) reporting of regulatory violations by all individuals as well as organizations. In practice, organizations use this protocol more frequently than individuals. Therefore, this approach is perceived by the industry as primarily an organization-level disclosure rather than individual-level disclosure. The current ASAP program is focused on the individual making the self-disclosure, providing specific legal protection to that individual as well as supporting a collaborative relationship between the FAA, the Company, and the Labor Union.

Philosophically, there seemed to be an agreement between the FAA and the maintenance community that the mechanic who actually commits the error holds key information that is essential to the

development of a truly comprehensive solution. Such an agreement is supported by extensive research in the area of error causation (Battles, Kaplan, Van der Schaff, & Shea, 1998; Gambino & Mallon, 1999; Van der Schaff, 1991 cited by Harper & Helmreich, 2003). The erring mechanic has no incentive (other than an ethical obligation) to disclose his/her error unless there was an effective non-punitive process in place.

Reporting Behavior in Maintenance

It is evident from the exponential rise in the number of ASRS reports filed by mechanics since 1996 that mechanics are willing to report their errors (Patankar & Taylor, 2001). In a recent study of reporting behaviors among 178 maintenance personnel in Australia, Fogarty (2003) reported that organizational factors/culture had a strong influence on the individuals' willingness to report maintenance errors. Fogarty concluded "employees were more likely to report mistakes in situations where management is communicative, open, and committed to safety values." In a similar study, Harper and Helmreich (2003), listed the following as factors that may influence an individual's willingness to report their own error: (a) mandatory versus voluntary system, (b) reporter protection, (c) ability to affect change, (d) fear of litigation and disciplinary action, (e) attitude toward the use of current reporting systems, (f) ease of use of the new/proposed system, (g) personal responsibility to address changes, and (i) management's endorsement of the new/proposed reporting system. The Maintenance ASAP Questionnaire (MAQ) developed for this study provides an opportunity to specifically test the mechanics' willingness to report their own errors—among a national sample.

Trust Between Mechanics and Managers

Interpersonal trust between mechanics and managers has been studied and extensively reported by Taylor and Christensen (1998) and Patankar and Taylor (2004). Based on these studies, it is known that there is a wide variation in such trust among the various maintenance organizations—interpersonal trust tends to be higher in smaller organizations and military units and lower among larger organizations—the range of trust values seem to indicate that up to a third of the mechanics don't tend to trust that their supervisors will act in the interest of safety.

Considering that interpersonal trust among mechanics, managers, and FAA inspectors was mentioned repeatedly during the focus-group discussions conducted earlier (Patankar & Driscoll, 2004), it was essential to include questionnaire items

associated with the “supervisor trust and safety” scale (Taylor & Thomas, 2003) in the MAQ.

METHODOLOGY

The Maintenance ASAP Questionnaire (MAQ) was developed from the responses to a series of focus-group discussions held at three organizations with ASAP programs and three organizations without ASAP programs (cf. Patankar & Driscoll, 2004).

A total of 104 items were created and the participants were asked to rate their level of agreement with each item on a 5-point Likert-type scale: 0= not applicable or don’t know, 1= strongly disagree, 2= disagree, 3= neutral, 4= agree, and 5= strongly agree.

All participants were expected to respond to the first 20 items; only the FAA inspectors were expected to respond to items 21-36; only the employees of organizations with ASAP programs were expected to respond to items 37-68, and only the employees of organizations without ASAP programs were expected to respond to items 69-104. Considering the similarities and differences in the items that each group (FAA inspectors, employees from organizations with ASAP programs, and those without ASAP programs) responded to, some common and some different scales emerged through subsequent factor analysis.

Currently, there are no known means to clearly establish, or even estimate, the number of FAA certificated mechanics and managers working for air carriers or approved repair stations. As of January 1, 2004, the FAA’s airman certificate database contained 230,880 Aircraft Mechanic certificate holders; however, there is no way of determining exactly how many of them are actively working as mechanics. Assuming that over 100,000 Aircraft Mechanic certificate holders are likely to be working for either an air carrier or a repair station, a minimum of 400 responses were required—“beyond a certain point (N=5,000), the population size is almost irrelevant and a sample size of 400 will be adequate” (Gay & Airasian, 2003, p. 113). As with any other survey, another obvious limitation of this study is that survey respondents tend to “self-select”—people who are interested in responding are likely to respond; to what extent the sample size is actually representative of the total population continues to be a matter of debate. Nonetheless, every effort was made to reach a diverse, and fully representative, population.

In order to minimize the perception among the participants that this study is either a “company survey,” a “union survey,” or an “FAA survey,” the FAA’s Airman Certificate database (publicly available for download from the FAA’s website) was used to construct a stratified sample consisting of

randomly selected participants from each state in the country. The total population of FAA certificated mechanics was sorted by states and ten times the required sample size was selected. For example, the state of Alabama has 3,468 FAA-certificated aircraft mechanics with A&P ratings. According to Gay and Airasian (2003, p. 113), a sample of 240 responses would be its statistically adequate representation. In order to maximize the probability of receiving 240 responses, 2,400 subjects were selected from the state of Alabama. In total, approximately 83,000 questionnaires were mailed out nationwide. All questionnaires were mailed to the participants’ home addresses and they were provided with a reply-paid envelope to return the questionnaires directly to Saint Louis University.

RESULTS

A total of 5,022 responses, from all fifty states, were received: 1,548 of the respondents were from organizations with ASAP programs, 2,920 respondents were from organizations without ASAP programs, and 124 respondents were FAA inspectors; 430 respondents did not know whether or not their organization had an ASAP program.

Overall Comparison (All respondents)

A factor analysis of the first twenty items on the MAQ resulted two scales: *willingness to report errors* and *supervisor trust and safety*. On the overall willingness to report one’s errors, there was no statistically significant difference between companies with ASAP programs and those without ASAP programs. Significance tested was at 0.05 level and the Cronbach’s alpha for this scale was 0.60.

On the supervisor trust and safety scale, employees from organizations with ASAP programs tend to trust their supervisors significantly more than those from organizations without an ASAP program ($p < 0.01$). Cronbach’s alpha for this scale was 0.79.

Overall, we see that maintenance personnel are quite willing to report their errors; regardless of whether or not they have an ASAP program. However, when there is an ASAP program, there is a higher level of trust in the management—trust that the management will act on safety suggestions.

FAA Inspectors Only

Analysis of the items posed to FAA inspectors revealed two new scales, in addition to the ones described earlier: perceived importance of ASAP programs (Cronbach’s alpha = 0.92) and perceived effects of ASAP programs on enforcement abilities (Cronbach’s alpha = 0.84).

About 40% of the FAA inspectors think that ASAP programs are important; another 40% are

somewhat undecided—perhaps, this population could be convinced of the advantages of ASAP programs if better training materials were to be made available. Now may be a great “window of opportunity” to shift the perception about ASAP programs from neutral to positive.

High scores on the perceived effects scale would have indicated that the FAA inspectors have resources to support local ASAP programs, they are willing to let a mechanic learn from his/her errors without resorting to punitive actions, they would not necessarily write fewer violations because of the ASAP program, and they generally don’t believe that their enforcement capabilities are compromised. However, most respondents scored low in this scale.

ASAP programs represent a fundamental shift in the way FAA administers safety and compliance. About 47% of the respondents to the perceived effects scale are undecided and need to be better convinced of the effects of ASAP programs on their ability to issue enforcement actions as well as overall change in philosophy—from compliance to collaboration. Considering that the FAA wants to move toward a collaborative error reduction program, about 70% (includes the ones who indicated “neutral,” “disagree,” or “strongly disagree”) of its inspector workforce needs to be better informed regarding the philosophical change that needs to take place.

Participants from Organizations With ASAP Programs

Based on 1,548 responses in this category, four new scales (in addition to the willingness to report errors scale and the supervisor trust and safety scale) emerged: ASAP programs are likely to improve trust (Cronbach’s $\alpha = 0.90$), ASAP programs are being used at their maximum potential (Cronbach’s $\alpha = 0.86$), ASAP programs receive adequate support from supervisors and coworkers (Cronbach’s $\alpha = 0.85$), and ASAP results need to be communicated and the protocol needs to be standardized (Cronbach’s $\alpha = 0.71$).

About 54% of the respondents (from organizations with ASAP programs) think that ASAPs are likely to improve trust; about 14% of them don’t think that the ASAP programs would improve trust.

Just over 44% of the respondents don’t seem to think that their current ASAP programs are being utilized to their maximum potential; about 12% of the respondents do think that their programs are close to full potential. The factors that would lead to better utilization of the maintenance ASAP programs include leveraging with flight and dispatch programs as well as improved communication/dissemination of

success stories, and training regarding ASAP acceptance criteria.

Even at organizations with ASAP programs, about 32% of the employees believe that they don’t get enough support from their superiors—leads, supervisors, and senior management.

About 71% of the respondents believe that there needs to be a strong communication regarding ASAP programs, including publicizing of the success stories and standardizing the process further.

Participants from Organizations Without ASAP Programs

The next sample consisted of employees from organizations without ASAP programs ($n=2,920$). In addition to the two basic scales regarding willingness to report and supervisor trust, this sample also revealed the level of difficulty in buying into the benefits of ASAP programs (Cronbach’s $\alpha = 0.92$), reported on the state of organizational climate at the time of the survey (Cronbach’s $\alpha = 0.87$), and level of awareness about, or interest in, ASAP programs (Cronbach’s $\alpha = 0.74$).

About 50% of the respondents agree with the items that tend to value the benefits of an ASAP program. Therefore, one could say that even in companies without ASAP programs, many people believe that ASAP programs have some benefits to offer. Since these results are from organizations without ASAP programs, it is not surprising that about 36% of the respondents did not know about the benefits of ASAP programs, 12% of the respondents were neutral, and 2% of the respondents did not seem to value any benefits of the ASAP program.

About 59% of the respondents disagree that they have a poor organizational climate. Therefore, one could say that just because an organization does not have an ASAP program, it does not mean that the organization is suffering from a poor or unhealthy safety climate.

A low or negative response on the awareness scale indicates that the general awareness about ASAP programs is low among these respondents. About 42% of the respondents disagree that they have a high level of general awareness about ASAP programs and that they have taken the effort to either review their own company’s pilot/dispatch ASAP program or have visited other company’s programs. If those who clearly indicated that they either did not know about the subject or that they thought that the questionnaire item was not applicable to them are combined, over 92% of the respondents (again, these respondents are from organizations that do not have ASAP programs) do not have a high level of awareness about ASAP programs.

DISCUSSION

Generally, there seems to be a high willingness to report errors; yet, there is also an overwhelming degree of mystery about ASAP programs. This is a great opportunity for the aviation maintenance industry to publicize the benefits of ASAP programs through dissemination of success stories and frequent open discussions with the mechanics from various line and base maintenance stations.

Since this survey indicates that organizations with ASAP programs have a higher degree of interpersonal trust and the overall maintenance community is struggling to raise this trust level in order to improve both quality of maintenance as well as the overall work environment, it would be worthwhile for companies to use collaborative programs such as ASAP to improve trust between mechanics, managers, and FAA inspectors.

Another important point to consider is that a substantial proportion of the respondents are “on the fence” regarding the benefits of ASAP programs—if such programs are to gain further momentum and achieve their full potential, this undecided population will need further proof and convincing that the ASAP programs are actually producing systemic changes without penalizing the reporters. Open meetings, traveling “road shows,” periodic status updates, dissemination of success stories through newsletters, and an overall advertising of the various changes effected by ASAP programs could lead to increased awareness of its benefits as well as increased trust in the process.

Also, field observations, focus-group discussions, and analysis of select MAQ items tend to indicate that there is limited leveraging of ASAP data across flight, maintenance, and dispatch groups. Any attempts to foster such tripartite leveraging could lead to novel, synergistic advances in safety and quality.

CONCLUSIONS

In conclusion, the factors that tend contribute toward a successful ASAP program in aviation maintenance organizations are as follows:

- There is a significantly higher level of trust between mechanics and their supervisors
- End-users perceive ASAP programs to be very valuable in improving the overall safety of the industry
- Good communication about the ASAP program and a standardized or a well-understood report handling process exists

Factors that contribute toward the failure of an ASAP program in aviation maintenance organizations are as follows:

- There is a significantly lower levels of trust between mechanics and their supervisors
- End-users don't seem to see a significant benefit in having an ASAP program—it is likely that they are satisfied with their internal error/hazard reporting program
- There is a severe lack of awareness about ASAP programs

Ultimately, one could combine the above success/failure factors into two key themes:

- Level of employee-management-FAA trust
- Level of awareness about ASAP programs

Focus group discussions on this topic indicate that this trust is influenced by experience with internal safety programs, success with past safety programs, and general labor-management relationship. Awareness, on the other hand, is a matter of consistent and concerted advertising of the effects of ASAP programs as well as soliciting of feedback to improve the program.

REFERENCES

- FAA (1998) (U. S. Federal Aviation Administration). Advisory Circular # 00-58 Voluntary Disclosure Reporting Program. Retrieved on January 13, 2004 from <http://www1.faa.gov/avr/afs/acs/00-58.txt>
- Fogarty, G. (2003). Errors, violations, and reporting behavior in aviation maintenance. In R. Jensen (Ed.) Proceedings of the Twelfth International Symposium on Aviation Psychology (pp.385-390), Dayton, OH.
- Gay, L., & Airasian, P. (2003). Educational research: Competencies for analysis and applications. Englewood, NJ: Prentice Hall.
- Harper, M., & Helmreich, R. (2003). Creating and maintaining a reporting culture. . In R. Jensen (Ed.) Proceedings of the Twelfth International Symposium on Aviation Psychology (pp.496-501), Dayton, OH.
- Marx, D. A., & Graeber, R. C. (1994). Human error in aircraft maintenance. In N. Johnston, N. McDonald, & R. Fuller (Eds.), Aviation Psychology in Practice (pp. 87-104). Aldershot, UK: Avebury.
- Patankar, M., & Driscoll, D. (2004). Preliminary analysis of Aviation Safety Action Programs in aviation maintenance. In M. Patankar (Ed.) Proceedings of the First Safety Across High-Consequence Industries Conference, [CD-ROM], St. Louis, MO, pp. 97-102.
- Taylor, J. & Thomas, R. (2003). Toward measuring safety culture in aviation maintenance: The structure of trust and professionalism. *International Journal of Aviation Psychology*. Vol 13(4), 321-343. Lawrence Erlbaum, US.
- Patankar and Taylor (2004). Risk management and error reduction in aviation maintenance. Adershot, UK: Ashgate Publishing, Ltd.
- Taylor, J. & Christensen, T. (1998). Airline Maintenance Resource Management. Warrendale, PA: Society of Automotive Engineers.

COMPUTER AND BROADBAND TECHNOLOGY IN AIRCRAFT LINE MAINTENANCE: A TASK ANALYSIS AND QUESTIONNAIRE

Stephen M. Casner, NASA Ames Research Center
Charles M. Encinas, Continental Airlines
Antonio Puentes, San Jose State University

ABSTRACT

The line maintenance work process was documented at two major air carrier facilities. This analysis shows how computer and broadband technology is used in most every phase of the line maintenance process with one important exception: maintenance technicians at neither carrier used technology on the ramp when performing maintenance on aircraft. We devised a questionnaire to query technicians' attitudes about their work process and whether or not their work could be improved by the use of technology. Questionnaire responses suggest three specific ways in which computer and broadband technology might support the performance of maintenance tasks.

TASK ANALYSIS

In previous studies, interviews have been used to determine those aspects of the maintenance process in which computer and broadband technology are presently being used, and what impact these technologies have on the work process [Casner and Puentes, 2003; Iyengar et al, 2004]. One limitation of the interview methodology is that it is generally difficult for technicians to recall all phases, aspects, and details of their work process during a brief conversation with an interviewer.

In this study, we documented the work process at line maintenance facilities of two major air carriers by observing technicians as they worked during regular work shifts. These observations were used to create a detailed description of the steps required to plan for and execute line maintenance required for the typical inbound flight. For each step in the work process, we noted all computer and broadband technologies that were used.

The task analysis was created during a series of four visits to two different air carrier maintenance facilities. During each visit, we were permitted to follow a single line maintenance crew consisting of two maintenance technicians for the duration of their work shift. We were allowed to follow the crews wherever they went and ask questions at any time.

Arriving On Shift

Maintenance technicians arrive to work, gather tools and safety gear, and learn two important things that will chart the course of their work shift: (1) who their partner will be for the shift; and (2) which scheduled flights they are assigned to meet, along with the gates at which the flights will arrive. Partner and flight assignments are made by the lead technician, prior to the arrival of the other technicians.

Planning and Preparation

Once the technicians know to which flights they have been assigned, they can begin preparing for the arrival of each aircraft. The maintenance tasks that must be performed on each aircraft can come from a variety of sources. Before the arrival of each flight, technicians become aware of two types of required maintenance tasks.

Routine Checks

A routine check is a basic walk-around of an aircraft that is required for most every incoming flight. Technicians must follow a published procedure for this check. These procedures are found in the maintenance manual for each type of aircraft.

Technology In Use: Maintenance manuals are stored in electronic format and are available on any of several computer workstations found in the maintenance office. Technicians must look up the procedures in the computer then print out a copy to be carried out to the aircraft when the procedure is performed. Casner and Puentes (2003) found electronic documentation systems to be used at every maintenance facility surveyed. Electronic documentation systems appear to solve a number of problems suffered by traditional paper manuals. First, technicians no longer have to wait to use a limited number of copies of any one manual: any manual can be accessed by any number of technicians at once when a computer terminal is available. Second, electronic documentation allows technicians to access any number of different manuals at a single computer workstation. Third, manuals can be revised electronically in a matter of minutes, and is a less error-prone process. For these reasons and others, electronic documentation systems enjoy widespread acceptance among technicians and managers.

Assigned Maintenance Tasks

All maintenance activities that occur across the company are monitored and managed by a central maintenance organization that manages all aircraft in the fleet as they travel from airport to airport. This organization can assign maintenance tasks to be performed upon arrival at particular airports. The local technicians are responsible for completing the assigned task(s) when the flight arrives.

An assigned maintenance task usually represents a known problem with the aircraft that might require troubleshooting, parts replacement, and/or significant work. Upon learning about assigned maintenance tasks, technicians can prepare in advance for the arrival of the aircraft. Technicians must gather three important resources that will allow them to perform the maintenance task: (1) documentation; (2) parts; and (3) tools. Documentation includes the relevant pages of the maintenance manuals and parts catalogs relevant to the task. These pages must be carried out to the aircraft when the work is performed. To gather parts, the technician must learn the relevant part numbers, determine if the parts are available from the company parts inventory, then retrieve the parts. Part numbers can be identified from the illustrated parts catalog (IPC). Typically, if maintenance control assigns a maintenance task, they will arrange to have needed parts delivered to the maintenance facility in advance. Once the technician verifies that the needed parts are available, s/he must walk over to the parts storage and retrieve the parts. If special tooling is needed to perform the maintenance task, these tools must be retrieved from a tool storage facility also located on the airport ramp. Once all documentation, parts, and tools are gathered, they can be loaded on a cart that will be driven out to meet the aircraft when it arrives.

Technology In Use: Technicians learn about assigned maintenance tasks by using a maintenance management software system installed on the computer workstations in the maintenance office. When a technician is assigned to work a scheduled flight, the technician must enter the flight number into the maintenance management software system. The system will then display all maintenance tasks that have been assigned by maintenance control.

Documentation such as maintenance manuals and illustrated parts catalogs are stored electronically. Pages from these manuals can be printed. Another software system allows technicians to look up part numbers and quickly determine whether or not a part is available at the facility. The system also tells

technicians where the part is located in the parts inventory.

Arrival At The Gate

Just prior to the scheduled time of arrival for each aircraft, technicians drive the cart equipped with documentation, parts, and tools out to the gate to which the incoming aircraft has been assigned.

Maintenance Problems Reported By The Flight Crew

After completing some of the routine check, technicians typically enter the cockpit to greet the flight crew. As the crew finishes their duties, technicians query the flight crew about any maintenance issues that arose during the flight. Although the crew is required to fill out a maintenance sheet for each maintenance discrepancy, the conversations between technicians and flight allow much more information to be exchanged than what is typically written on a maintenance sheet. Specifically, the expert technicians are able to ask questions of the flight crew to clarify or give more details about maintenance problems.

Technology In Use: As an incoming flight comes within radio communication range of the airport, the flight crew can call in other maintenance problems they have experienced. These in-range calls are designed to give maintenance technicians extra time to prepare for unexpected maintenance tasks. These calls are made via VHF radio transmissions from the cockpit to the lead technician in the maintenance office. Once received in the maintenance office, this information is passed on to the technicians that have been assigned to meet the aircraft. Upon learning about these maintenance problems, technicians must quickly go through the same preparatory routine they have done for any assigned maintenance tasks: gathering documentation, parts, and tools.

Maintenance Problems Discovered During Routine Checks

Technicians then complete the walk-around of the aircraft. This routine check represents another source of maintenance discrepancies and tasks: those discovered during the routine check.

Technology In Use: Many airplanes contain onboard diagnostic computers that can automatically detect faults during the flight. Technicians access this information after the crew leaves and they complete their cockpit checks. This represents another class of maintenance problems: those detected by the computer but that were unknown to the flight crew.

At this point, the technicians now know about all of the maintenance problems that they will have to deal with during the airplane's stay at the airport. To recap, these problems have come from four different sources:

1. Tasks assigned by maintenance control
2. Problems discovered during routine checks
3. Problems reported by the flight crew (in-range or on the ground)
4. Problems reported by the airplane's on-board diagnostic computers

Troubleshooting and Solving Maintenance Problems

Confronted with a list of maintenance tasks, technicians have one overriding goal: to do everything possible to ensure that the aircraft is able to depart on schedule. There are two basic ways to address each maintenance problem: (1) deferring the problem; or (2) resolving the problem.

Deferring Maintenance Problems

Many types of maintenance problems can be deferred for specified periods of time. This is the most desirable option for problems other than those that can be resolved quickly. Deferral allows the aircraft to depart on schedule, and also allows maintenance control to assume responsibility for the maintenance problem. Recall that maintenance control commands all of the technical resources of the entire company. A deferral allows maintenance control to determine which of the aircraft's upcoming stops would be best suited for a particular type of maintenance problem. Maintenance control can choose an airport that has the most appropriate technicians, arrange to have needed parts or tools made available, and choose the stop that offers technicians the most time to work on the problem.

To defer a maintenance problem, the crew must determine whether or not the problem is legally deferrable. A document called a minimum equipment list (MEL) records the list of parts that can be inoperative for any aircraft. If a maintenance problem amounts to an inoperative part, and that part can be found on the minimum equipment list (MEL), technicians can legally defer the problem and the aircraft can depart on schedule.

Resolving Maintenance Problems

There are two kinds of maintenance problems that are not deferred: (1) those that are not deferrable according to the minimum equipment list (MEL); and (2) those that have previously been deferred, and can be deferred no longer. Tasks that are assigned by maintenance control are typically of the second

variety: tasks that were deferred by technicians during previous stops.

Resolving a maintenance problem represents the real work of the maintenance technician. Technicians must now use their knowledge and skills to isolate and remedy each problem. Technicians have a variety of resources available to them when resolving a maintenance problem.

Documentation Resources

Technicians have several documentation resources available to them when resolving a maintenance problem. A fault isolation manual (FIM) prescribes a series of steps to be used when troubleshooting a problem. The steps in the FIM involve replacing parts, one after another, until a faulty part is found and replaced and the system functions normally again. When replacing each part, technicians must return to the maintenance office, look up the part number, determine if the part is in stock, then return to the airplane to replace the part. If this part turns out not to be the defective part, these steps must be repeated. In many cases, if a part is replaced and it does not result in a fix, that part remains in the aircraft and the old part is retired, or must be recertified before it can be used again in another aircraft.

In the case that the procedure in the fault isolation manual does not result in resolution of the problem, technicians must resort to other troubleshooting resources. Circuit diagrams allow technicians to trace through electrical circuits when troubleshooting.

Technology In Use: Documentation is stored electronically and available at the computer workstations in the maintenance office. At one of the maintenance facilities, laptop computers were available. These computers allowed technicians to access documentation and make entries into the maintenance management software remotely. We did not observe a single instance of a technician using these computers.

Other Technicians

A variety of human resources are available to technicians when working on a problem. Technicians can consult with other technicians working on other aircraft on the ramp. Technicians can call the lead technician and ask for assistance. Maintenance control offers technical assistance on any maintenance topic via telephone.

Technology In Use: Technicians often use company radios or personal cell phones to talk when away from each other.

Wrapping Up

After work is completed at the aircraft, technicians return to the maintenance office and make entries in the maintenance records for the aircraft. In the case of a deferral, the technician records the deferral. In the case that a problem is resolved, the technician records all of the maintenance actions that were taken, and certifies that the aircraft can be returned to service. In the case that a problem is neither deferred nor resolved, the aircraft must be grounded.

Technology In Use: Technicians make entries into maintenance records using the same maintenance management software. This system makes the maintenance just performed available to technicians and managers across the company.

QUESTIONNAIRE

The analysis above describes how broadband and computer technologies are used in most phases of the line maintenance process except for one: troubleshooting and solving maintenance problems. To investigate the reasons for why technology is not used in this central part of the maintenance process, we developed a paper and pencil questionnaire.

Questionnaire items were designed to explore three questions raised by our task analysis:

1. Do technicians feel that current documentation systems well support the performance of maintenance tasks?
2. How much importance do technicians place on each others' expertise, and how well does current technology support the sharing of expertise?
3. Are technicians open to the idea of using computer and broadband technology while working out on the ramp?

Participants

Sixty-eight maintenance technicians participated in the study on a voluntary basis. Technicians who completed the questionnaire were given a NASA t-shirt as compensation.

Apparatus

The questionnaire contained thirty-four questions and covered both sides of a single sheet of paper. The questionnaire items, listed below, were designed to probe technicians' opinions about the resources they currently have available to them when resolving maintenance problems, and what resources they might find desirable in the future. Since our focus was on technological resources, our questionnaire also queried

technicians about their experience with computers and broadband technology.

Each questionnaire item made a statement about resources that might be used during line maintenance, and asked participants to agree or disagree with the statement using a five-element Likert-type scale.

Questionnaire Items

1. When troubleshooting a problem, the fault isolation manual (FIM) usually provides everything I need
2. The FIM is usually the best way for an experienced technician to troubleshoot a problem
3. The FIM is usually the best way for an inexperienced technician to troubleshoot a problem
4. I often use other sources of information (i.e., wiring diagrams) in addition to the FIM
5. I always follow the steps in the FIM exactly as written
6. I often consult with other technicians
7. I often consult with maintenance control
8. I often consult with the lead or supervisor
9. I can often provide information to other technicians that can help them troubleshoot a problem
10. Other technicians often provide me with useful information
11. Someone on my shift always knows the answer to my question
12. Experienced technicians often provide better information than the manuals
13. Different technicians excel in different areas of expertise
14. Technicians should learn to find the information rather than asking me for it
15. Technicians can learn a lot just by talking to each other
16. I would rather use the manuals than ask another technician
17. Communication between technicians at our facility is adequate
18. We should have a better way for technicians to talk to each other at our facility
19. I often use company radios to talk to other technicians on the ramp
20. If other technicians have already solved a difficult maintenance problem, I'd like to have their notes in front of me when I'm dealing with that same problem
21. It would be nice to have some kind of searchable database of difficult maintenance problems

22. This searchable database should allow technicians to enter any relevant notes about procedure, tooling, etc.
23. A searchable database should allow us to use any keywords, like a web browser
24. Searching maintenance histories using ATA codes alone is too limiting
25. I would be willing to submit information about difficult maintenance problems to this database
26. I think most other mechanics would be willing to submit information to this database
27. Finding information in the computerized maintenance manuals is relatively quick and easy
28. I wish the manual were more easy to search or use
29. I wish there was a way to more quickly access needed information when I'm out at the aircraft
30. There should be an easier way for me to access frequently-used information like tire pressures and torque values (e.g., a "quick reference")
31. Having a quick reference for frequently-used information would increase my productivity
32. I would use a PDA (e.g., Palm Pilot) to access maintenance information at the aircraft
33. I would use a laptop computer to access maintenance information at the aircraft
34. Using computer equipment of any kind or size at the aircraft is cumbersome

Procedure

Questionnaires were distributed to line maintenance technicians working at three different facilities operated by the same airline company. Questionnaires were handed to line maintenance technicians at the beginning of several work shifts by the lead technician who served as supervisor for the shift.

Results and Discussion

Figure 1 shows the mean and standard deviation for responses given to each questionnaire item. These statistics were derived by numerically coding the five-element Likert-type scale used to elicit responses from participants. Scores of 1 through 5 were assigned to responses of Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree, respectively.

Do technicians feel that current documentation systems support the performance of maintenance tasks?

In response to item 4, technicians agreed that they used other documentation materials besides the fault

isolation manual (FIM) (e.g., wiring diagrams) when troubleshooting problems [4.24 (0.8)]. Technicians provided the strongest response to item 29, indicating that they wanted a means of more quickly accessing these resources when out at an aircraft [4.33 (0.61)]. Item 30 indicated that technicians wanted an easier way of accessing frequently-used information, while item 31 showed that technicians believed that having this quick access would increase their productivity.

ANALYSIS: Although electronic documentation enjoys good acceptance overall [Casner and Puentes, 2003] and seems to well support the maintenance planning process, the current use of computer workstations in the maintenance office may not well support the sometimes iterative process of troubleshooting and solving problems. Having to walk back and forth between office and aircraft seems to be a burden for technicians. Questionnaire responses suggest the need for a means of remotely accessing electronic documents while working on the ramp.

In addition, technicians expressed the need to have some information items, found within a document, to be more readily accessible. In our earlier interviews [Casner and Puentes, 2003], technicians often complained about having to access manual pages in the computer for numbers that they use everyday.

How much importance do technicians place on each others' expertise, and how well does current technology support the sharing of expertise?

Item 6 was the most direct (I often consult with other technicians) and received an average response of 3.96 (0.76). Responses to items 9 and 10 suggest that technicians generally agree that they have valuable information to share with other technicians, and that they benefit from information provided by other technicians. Technicians strongly responded to item 13, that different technicians excel in different areas of expertise [4.13 (0.8)], and to item 15: that technicians can learn a lot from talking to one another [3.91 (0.64)].

ANALYSIS: Responses to these questionnaire items indicate that communication among technicians is a core part of the maintenance crews' problem-solving capability. It is important to note that sharing of expertise can happen on two different time-scales. As indicated by item 19, technicians sometime use radios to talk to each other while working on the ramp. We saw many instances of cell phone use for the same purpose. This allows technicians to communicate with one another on a minute-by-minute basis: asking simple questions and coordinating movements while

out on the ramp. But sharing of expertise can also happen on a wider time scale: the idea of passing on information gained through experience to crews working future shifts or crews at different locations. Items 21, 22, and 23 addressed this idea by probing technicians' interest in having some sort of searchable database that provides case-specific information about

previous difficult maintenance problems. Technicians generally agreed that they would like to have such a system [4.0 (0.85)], that this database should allow technicians to enter relevant notes about each case [4.01 (0.66)], and that the database should be searchable using keywords, like a web browser [4.1 (0.63)].

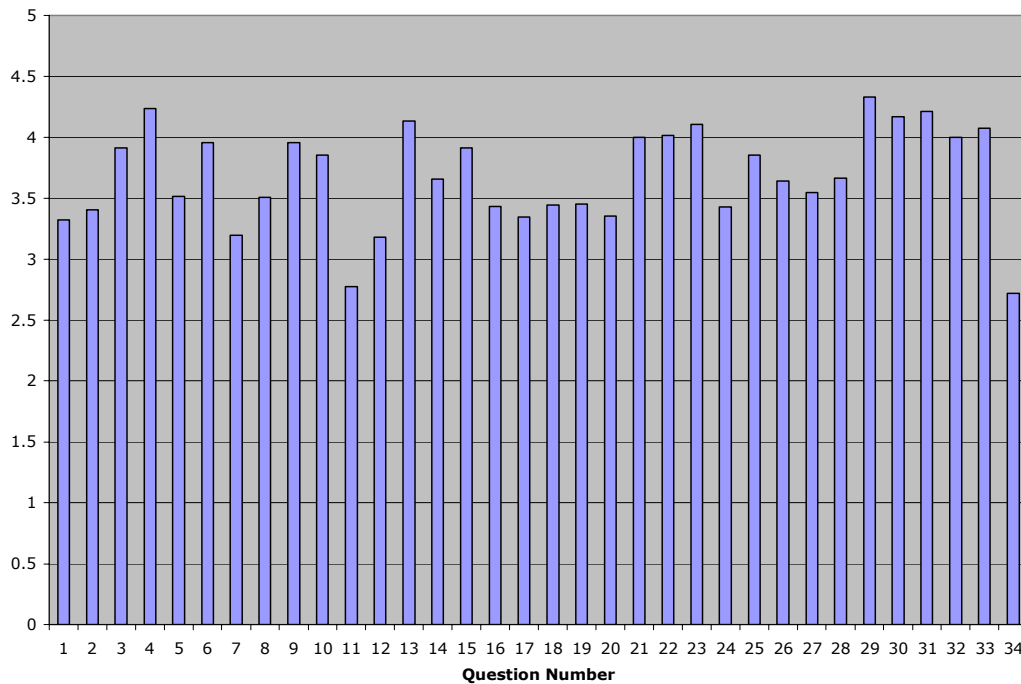


Figure 1: Average responses to questionnaire items

Are technicians open to the idea of using computer and broadband technology while working out on the ramp?

Questions about technicians' willingness to use portable computer technology on the ramp was motivated by our observation that technicians had laptop computers available at one facility, but did not use them.

Item 33 directly asked technicians if they would be willing to use laptop computers on the ramp. This questionnaire item yielded a response of 4.07(0.8), a result contrary to what we observed.

Item 32 asked technicians if they would consider using a PDA device while working at the airplane. The average response was 4.0 (Agree), and there was no difference between technicians who had [3.92 (0.93)]

and had not [4.05 (0.82)] previously used PDA devices.

ANALYSIS: Despite the agreement that portable computers would be useful, our task analysis revealed no use of such devices, even though wireless laptop computers were available to technicians at one of the facilities we visited. We interpret this lack of use as an indication that the portable computers fail to offer the functionality, usability, or reliability that technicians seek in such a device.

CONCLUSION

To summarize, we draw three conclusions from what we observed during the task analysis and the responses to the questionnaire.

First, although current electronic documentation systems solve many problems suffered by traditional paper manuals, it seems that current documentation

systems could evolve in specific ways. First, documentation should be accessible by means other than the computer workstations located in the maintenance office. Technicians expressed a need to access manuals while working out on the ramp. Second, it does not appear that current documentation systems fully exploit the advantages of the digital medium. In many cases, manuals are simply digitized versions of a paper manual. Technicians expressed a need to more quickly search for frequently-used items in the manuals. This suggests document search functionality that goes beyond the typical index and table of contents. This functionality might even extend to the idea of a document that dynamically reorganizes itself depending on how the document is searched and used over time.

Second, observations and questionnaire responses strongly suggested that technicians rely on each other when solving maintenance problems. Our observations showed that technicians frequently use cell phones and company radios to talk with one another while working on the ramp. Technicians also make use of a telephone help system that allows them to call the company's maintenance control facility to get advice from other technicians who specialize in particular areas. Questionnaire responses indicated that technicians would like one additional resource: a database system that allowed them to access notes left by other technicians from previous maintenance problems. Such a system would provide yet another means of sharing expertise between technicians.

Questionnaire responses indicated that technicians would not only use such a system, but also be willing to submit their own notes to such a system.

Third, with regard to using portable computers while working on aircraft, questionnaire responses contradicted the behavior we observed during the task analysis. Technicians claimed they would be willing to use portable computer, yet did not use them in practice when they were made available. This suggests that the design of presently-available portable computers does not match what technicians are looking for in such a device.

Future Work

With these three conclusions in mind, we have begun prototyping a hardware/software tool (illustrated in Figure 2). The purpose of this tool will be to explore the idea of providing technicians with the capabilities that they appear to need and have claimed to want. This tool, implemented on a PDA device, will evaluate the feasibility of offering technicians three capabilities:

1. Portable access to existing maintenance documentation;
2. A means of more quickly accessing frequently-used documentation items;
3. A means of searching and contributing to an archival database of previous maintenance cases;

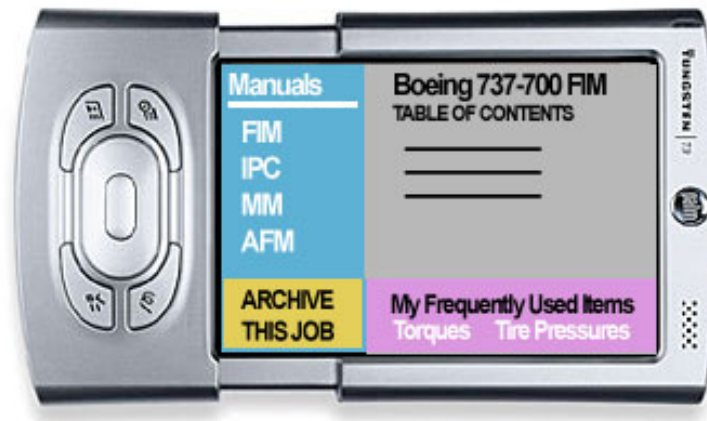


Figure 2: Prototype portable tool

REFERENCES

Casner, S. M., and Puentes, A. (2003). Computer and broadband technology in the aviation maintenance workplace. Human Factors Aviation Maintenance Program Review FY03.

Iyengar, N., Kapoor, K., Greenstein, J. S, and Gramopadhye, A. K. (2004). Data gathering

methodologies to identify impact variables in aviation maintenance. In Proceedings of the Industrial Engineering Research Conference.

Fatigue and Aircraft Inspection 1: Model and Simulation

Colin G. Drury, Monique Saran and John L. Schultz
University at Buffalo, Department of Industrial Engineering
438 Bell Hall, Amherst, NY 14260
drury@buffalo.edu

Abstract

This paper examines issues of fatigue in inspection by using an established function analysis of inspection to show its characteristics, and then proposing a four-level classification of temporal effects to help future applications. This classification divides the temporal effects into four components: weekly, daily, hourly, and minute time scales. The analysis presented here formed the basis for the design of a simulator for Fluorescent Penetrant Inspection of engine blades to be using in experimental studies of temporal factors in aircraft inspection. Initial Results from 8 participants are presented.

Introduction

Failures of both airframe inspection and engine inspection have highlighted the potential impact of human limitations on inspection system performance. Accidents that have occurred due to engine inspection failure include the Sioux City and Pensacola accidents. The 1989 Sioux City crash was the result of inspection not finding a crack in an engine disk. Remnants of fluorescent penetrant were found in the crack after the crash. These remnants helped to determine that the crack was large enough to be seen when the inspection occurred but why it was missed is not known. The 1996 Pensacola crash was due to a fan hub in the left engine having an undetected crack. Both of these crashes could have been prevented if the cracks had been located during inspection. In a 1998 incident to an Aloha Boeing 737 aircraft, evidence was found of multiple site fatigue damage leading to structural failure. The resulting National Transportation Safety Board investigation report issued in 1989 attributed the incident to the failure of the operators' maintenance program to detect corrosion damage. A number of visual and Non-Destructive Inspection (NDI) techniques require the inspector to work continuously on repetitive tasks for extended periods. Examples are fluorescent penetrant inspection of engine rotor blades, eddy current inspection of large batches of wheel bolts, and magnetic particle inspection of landing gear components. Such tasks typically occur on all shifts and can involve inspecting at low periods of the human circadian rhythm. Inspectors may be subject to the effects of cumulative fatigue from overtime and shift work.

In all of these inspection tasks, the *a priori* similarity to classical vigilance tasks suggests that performance (defect detection) may decrease with time spent inspecting. However, much skepticism exists regarding the relevance

of vigilance studies to the operational environment. In the case of aircraft inspection tasks, there is the added complication of the relevance of shift-work and circadian rhythm studies to these particular tasks. Thus, we have two issues:

1. Can we expect the findings from the vigilance literature to apply to aircraft inspection?
2. How well might the studies of circadian rhythms and cumulative fatigue from shift working apply to vigilance, and then to aircraft inspection?

Note that both of these issues concern the temporal effects of inspection work. This paper examines these issues by using an established function analysis of inspection to show its characteristics, and then proposing a four-level classification of temporal effects to guide future applications. Indeed, the analysis presented here will form the basis for the design of future experimental studies of temporal factors in aircraft inspection.

Analysis of Inspection Tasks in Aviation

To understand inspection, and to provide a link between inspection and the psychology / human factors literature, we use the generic functions which comprise all inspection tasks whether manual, automated or hybrid. We have recently undertaken a systematic analysis of all of the inspection techniques involved in NDI of aircraft (Drury, 2003), so far covering Fluorescent Penetrant Inspection (FPI), Visual inspection, Borescopes, Eddy Current and Ultrasonics. All were studied in aircraft maintenance settings to perform Hierarchical Task Analyses and thus derive a set of Good Practices related to human and system functioning. Each of these NDI techniques exhibited all of the generic functions, although some required much preparation prior to the actual inspection. Table 1 shows these functions, with the specific application to NDI in

aviation. The functions of search and decision are the most error-prone, although for much of inspection, especially NDI, setup can cause its own unique errors (Murgatroyd, Worrall and Waites, 1994). Search and decision have been the subjects of considerable mathematical modeling in the human factors community, with direct relevance to visual inspection.

In the visual aspects of inspection tasks, the inspector must move his/her eyes around the item to be inspected to ensure that any defect will eventually appear within an area around the line of sight in which it is possible to achieve detection. This area, called the visual lobe, varies in size depending upon target and background characteristics, illumination and the individual inspector's peripheral visual acuity. As successive fixations of the visual lobe on different points occur at about three per second, it is possible to determine how many fixations are required for complete coverage of the area to be searched. We have useful models of visual search applicable to inspection (Wolfe 1994; Drury and Hong 2000), but the point made here is that all inspection tasks in aviation do involve some search, in contrast to many laboratory vigilance tasks.

Decision-making is the second key function in inspection. This is where each indication is judged as being a defect or not a defect. An inspection decision can have four outcomes (Table 2). These outcomes have associated probabilities, for example, the probability of detection is the fraction of all defective items rejected by the inspector shown as p_2 in Table 2.

Decision of Inspector	True State of Indication	
	Non-defect	Defect
Accept, i.e. Call non-defect	Correct accept, p_1	Miss, $(1 - p_2)$
Reject, i.e. Call defect	False alarm, $(1 - p_1)$	Hit, p_2

Table 2. Four outcomes of inspection decisions

At this point, the obvious rational decision making models such as Signal Detection Theory are usually invoked to equate inspection to simple decisions. From the analysis in Table 1, it is clear that inspection is not merely the decision function. The use of models such as signal detection theory to apply to the *whole* inspection process is misleading in that it ignores the search function. For example if the search is poor, then many defects will not be located. At the overall level of the inspection task, this means that probability of detection (PoD) decreases, but this decrease has nothing to do with setting the wrong decision criteria. Even such devices as ROC curves should be applied only to the decision function of inspection, not

the overall process, unless search failure can be ruled out on logical grounds.

Temporal Aspects of Inspection

Temporal effects in the literature occur over four times scales:

1. Weeks, where the issues are shift work and cumulative fatigue from hours of work, sleep loss, days worked, overtime and shift work.
2. Days, where circadian rhythms are predominant, so that time of day is the main driver.
3. Hours, where the issues are times spent continuously on tasks, and the timing, nature and duration of rest periods
4. Minute, where the concern is sequential effects in repetitive tasks: does the detection of a defect on one item inspected affect the behavior or performance on subsequent items?

Each of these is reviewed in turn before examining in more detail their relevance to aircraft inspection. To help obtain background data on the hours of work and shift work patterns of NDI inspectors, a survey "Aircraft Maintenance Personnel Survey of Work Hours" was given to samples of NDI inspectors at several airlines. The survey, Folkard (2002), asks about hours of work, shift systems, breaks, vacation days and some symptoms of stress. Here we present simple summary statistics, from our first group of 40 NDI inspectors at two airlines. The sample was older and more experienced than typically found for AMTs. Comparing the age and experience distributions to the population demographics of Aviation Maintenance Technicians found in a national sample compiled by the Bureau of Labor Statistics (BLS, Washington, 1991), our sample was significantly older with a median age of 46.5 year versus a BLS median age of 36.2 years (Wilcoxon test, $t = 645$, $p < 0.001$). Our sample was also more experienced with a median of 24.0 years as an Aviation Maintenance Technician versus a BLS median of 9.4 years (Wilcoxon test, $t = 780$, $p < 0.001$). Selected questions on hours of work and rest are given in Table 3.

	Median	Minimum	Maximum
Hours of work per week	40	30	56
How long before a work break?	2.0	1.0	4.0
How many minutes does break last?	12.5	0	45
How many days annual leave?	31	11	40

Table 3. Sample work characteristics of NDI Inspectors

The temporal work characteristics appear about what would be expected, with 40-hour weeks, 2 hours between

breaks and 10-minute breaks. The relatively long vacation periods presumably arise from the high seniority typical of NDI inspectors, and confirmed here by the high age and experience statistics.

Weeks

The cumulative fatigue effects of shifts may span a period of a week or more. Fletcher and Dawson (2001) showed how fatigue builds up over the course of a week and its interactions with circadian variations. Their model was validated with a field study using OSPAT (Occupational Safety Performance Assessment Technology) performance tests and a VAS (Visual Analog Scale) measurement of alertness. French and Morris (2003) developed the FADE model that was validated using results from pattern recognition test from the NASA Space Cognitive Assessment Test (SCAT) battery and a divided attention version of the Maniken Task. Both models show the cumulative effects of shift work over a week and show circadian lows that occur daily.

Days

The daily variations in performance that an individual goes through are cyclic and predictable. The circadian rhythms or internal biological clock combined with environmental cues (*zeitgebers*) make people diurnal or active during the day. In general, humans show the same type of phasic behavior in performance as these biological rhythms, but there are individual differences in the timing of the onset of phases. Typically, people experience a circadian low, in measures such as body temperature, at approximately 0400 each day. Other variables relating to human bodily functions have been found to have lower values at night including heart rate, blood pressure and urinary excretion (Folkard 2002; Fletcher and Dawson, 2001). Studies of shift work contain strong evidence for circadian rhythm influence on performance decrements and contain recommendations for ameliorating performance decrements associated with circadian variations (Della Rocco, Comperatore, Caldwell, Cruz 2000; Fletcher and Dawson 2001; Folkard 2002). Vigilance effects (see **Hours**) appear quite sensitive to diurnal effects.

Hours

The vigilance decrement is a decline in performance that occurs along the hourly time scale. Typically, performance drops during the first 15 minutes on task and continues to decline until about 30 minutes into a task (Teichner, 1974).

Parasurman and Davies (1977) discussed vigilance in depth from a decision theory (SDT) approach and stated the decline in performance was based on the task

characteristics of successive vs. simultaneous and the event rate or the numbers of stimuli over time. Their taxonomy of vigilance showed that sensitivity decrement was related to these two factors. More recently, See, Howe, Warm and Dember (1995) conducted a meta analysis of the sensitivity decrement in vigilance and determined that these task characteristics are a large component of the vigilance decrement but that the sensory-cognitive component must be investigated as well. For aircraft inspection work this last distinction is not relevant, no targets are uniformly “sensory” in See et al, terminology.

Vigilance shares many characteristics of the inspection task such as rare signals, time on task, high memory load, and spatial and temporal uncertainty, but is different in other ways, as detailed later.

Minutes

Sequential effects are those found on time scales of seconds or minutes, and represent the influence of recent prior targets on subsequent performance. Tsao (1984) found that “following the detection of a faulty item, stopping time decreases for the second and third items, increases for the sixth and seventh items, and then levels off.” A re-analysis of the Panjwani and Drury (2003) data on rare-event inspection found a negligible sequential effect. There may be small sequential effects, but they are unlikely to influence the aircraft inspection task significantly due to the very low event rate for this task, and to their small absolute magnitude.

Relevance to Aircraft Inspection

From the site visits, the hours of work survey and Folkard’s study in the aviation maintenance industry, it does appear that temporal effects are likely in aircraft inspection tasks. Shift working is common, although most inspection in component shops is still on day shift. Both night shifts and changing shift schedules have been shown to reduce performance on tasks similar to inspection, e.g. vigilance tasks. While it is still not clear how closely vigilance mimics aviation inspection tasks, it is quite clear that vigilance tasks are particularly sensitive to the effects of circadian lows and cumulative fatigue from shift working. Thus, inspection tasks with vigilance-like characteristics are performed at times when decrements would be expected. The integrative models of Folkard (2002), Fletcher and Dawson (1998) and French and Morris (2003) all give sound advice on avoiding cumulative fatigue states. The typical work/rest schedule is 2 hours work followed by 10 minutes rest, which would again give cause for concern if vigilance tasks were indeed close mimics of inspection. The vigilance decrement literature shows performance declines over periods of less than one

hour for some types of vigilance task. Tasks particularly susceptible to decrements are those where there is no constantly available comparison standard, and where signals are rare, both characteristics of aircraft inspection. Other factors causing a vigilance decrement are less relevant: untrained personnel and symbolic stimuli. Overall, we can compare the attributes of classical vigilance tasks with those of aircraft inspection, as shown in Table 4.

As we move to the broader fields of temporal effects, such as circadian rhythms or shift work, we must not assume that vigilance findings hold. Indeed, a recent paper on time of day effects (Horowitz, Cade, Wolfe, and Czesler, 2003) found the usual effect of peaks and troughs of circadian rhythm on a vigilance task, but none on a simple search tasks performed at similar times.

FPI Simulation

The next step was to develop a methodology for measuring any temporal effects in inspection tasks relevant to aviation. The task had to be one that is performed repetitively over all shifts and where both misses and false alarms were possible. After examining a number of tasks, a FPI task was chosen, specifically of engine blades. We produced high quality photographs of all six faces of 63 different blades from a JT8-D engine at AANC. These were modified in Adobe Photoshop to match the penetrant colors under UV illumination, and realistic defects were added, with sizes based on PoD curve data. Penetrant “background” was added in realistic amounts, including covering all defects.

A program was written in Visual Basic to allow batches of blades to be inspected blade-by-blade and face-by-face. As each face is inspected, a swab tool can be used to remove penetrant “background” to determine whether or not it conceals a defect. A 4X magnifying glass tool can be used to enlarge portions of the blade for closer scrutiny. When a defect is detected, a dialog box allows participants to make a written report of its location and severity.

This simulation was first pre-tested on students, then taken to an airline partner for pre-tests by four FPI inspectors. They agreed that it was realistic and after input from them and FAA personnel, it was finalized for testing using experimental participants.

So far 8 participants from the local community have been run. In the first hour they are taking an average of 116s per blade to search for a defect, and finding 66% of the defects. The False Alarm rate is 8.5%. Participants either perform for 1 or 2 hours, with or without breaks each 20 minutes, and start at 0300 or 0900 to test circadian effects. So far there is typical learning in performance times, but no change in either p(hit) or p(False Alarm) over

time periods. No statistical analyses have been performed as the 8 participants represent only 10% of the complete 2⁵ experimental design.

References

- Bureau of Labor Statistics (1991). BLS, Washington.
- Della Rocco, P. S., Comperatore, C., Caldwell, L. and Cruz, C. E. (2000). *The Effects of Napping on Night Shift Performance*. FAA Tech Report.
- Drury, C. G. (2003). Human Factors and Inspection: Best Practices. *Proceedings of the 93rd Congress of the International Ergonomics Association Annual Conference*, Seoul, Korea.
- Drury, C. G. and Hong, S.-K. (2000). Generalizing from single target search to multiple target search. *Theoretical Issues in Ergonomics Science*. **1**(4), 303-314.
- Drury, C. G. and Prabhu, P. V. (1994). Human factors in test and inspection. In G. Salvendy and W. Karwowski (eds.), *Design of Work and Development of Personnel in Advanced Manufacturing*, J. Wiley, New York, 355-402.
- Fletcher, A. and Dawson, D. (1998). A Work-Related Fatigue Model Based on Hours-of-Work. *Managing fatigue in transportation: proceedings of the 3rd Fatigue in Transportation Conference*, Fremantle, Western Australia, 189-208.
- Fletcher, A. and Dawson, D. (2001). Field-based validations of a work-related fatigue model based on hours of work. *Transportation Research Part F* **4**, 75-88.
- Folkard, S. (2002). *Work Hours of Aircraft Maintenance Personnel*. Civil Aviation Authority, CAA Paper 2002/06.
- Folkard, S. and Monk, T. (1985). *Hours of Work, Temporal Factors in Work-Scheduling*, John Wiley and Sons, Chichester.
- French, J. and Morris, C. S. (2003). Modeling Fatigue Degraded Performance in Artificial Agents. *Proceeding of the Human Factors and Ergonomics Society 47th Annual Meeting*, 307-310.
- Horowitz, T. S., Cade, B. ., Wolfe, J. M. and Czesler, C. A. (2003). Searching night and day: a dissociation of effects of circadian phase and time awake on visual selective attention and vigilance. *Psychological Science*, **14**(6), 549-557.
- Huey, B. M. and Wickens, C. D. (1993). *Workload Transition: Implications for Individual and Team Performance*, The National Academy of Sciences.
- Murgatroyd, R. A., Worrall, G. M. and Waites, C. (1994). A Study of the Human Factors Influencing the Reliability of Aircraft Inspection, AEA/TSD/0173. Risley, AEA Technology.
- Panjawani, G. and Drury, C. G. (2003). Effective Interventions in Rare Event Inspection. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Denver, CO, 41-45.
- Parasuraman, R. and Davies, D.R. (1977). A taxonomic Analysis of Vigilance Performance. In R. R. Mackie (ed), *Vigilance, Theory, Operational Performance, and Physiological Correlates*, Plenum Press, New York.

- Teichner, W. H. (1974). The Detection of a Simple Visual Signal as a Function of Time of Watch. *Human Factors*, **16(4)**, 339-353.
- See, J., Howe, S. R., Warm, J. S. and Dember, W. N. (1995). Meta-Analysis of the Sensitivity Decrement in Vigilance. *Psychological Bulletin*, **117(2)**, 230-249.
- Tsao, Y.-C. and Wang, T.-G. (1984). *Inspectors' Stopping Policy After Fault Detection*. *Human Factors*, **26(6)**, 649-657.
- Wolfe J. M. (1994). Guided Search 2.0: A Revised Model of Visual Search. *Psychonomic Bulletin and Review*, **1(2)**, 202-238.

Function	Inspection Description
1. Initiate	All processes up to accessing the component. Get and read workcard. Assemble and calibrate required equipment. For FPI this includes part preparation steps.
2. Access	Locate and access inspection area. Be able to see the area to be inspected at a close enough level to ensure reliable detection. For component inspection, the parts are typically brought to the inspector rather than the inspector going to the airframe.
3. Search	Move field of view across component to ensure adequate coverage. Carefully scan field of view using a good strategy. Stop search if an indication is found.
4. Decision	Identify indication type. Compare indication to standards for that indication type.
5. Response	If indication confirmed, then record location and details. Complete paperwork procedures. Remove equipment and other job aids from work area and return to storage. If indication not confirmed, continue search (3).

Table 1. Generic function description and application to Non-Destructive Inspection

VIGILANCE TASK ATTRIBUTE	INSPECTION TASK ATTRIBUTE
Important Signals	Cracks or other defects that can have direct safety consequences.
Rare Signals	Defects can range from quite common, e.g. corrosive areas on older aircraft, to extremely rare (e.g. cracks in jet engine titanium hubs). Under most circumstances far less than 1 out of 10 inspected components will contain a reportable defect.
Low Signal Strength	Most defects are perceptually difficult to detect, often occurring within a background of non-defects, e.g. cracks among dirt marks and scratches.
Long Time on Task	Time on task can vary from a few minutes to about 2 hours without a break. Scheduled breaks are typically four 15-min breaks per shift, but many tasks are self-paced so that inspectors can break early or continue beyond scheduled time to complete an area or component.
High Memory Load	Prototypical defects are usually stored in the inspector's memory, rather than being presented as part of the task. Sometimes typical defects are illustrated on workcards, but workcards are often poorly integrated into the inspection task.
Low Observer Practice	Inspectors are highly skilled and practiced, after 3-10 years as an AMT before becoming an inspector. However, for some rare defects, even experienced inspectors may literally never have seen one in their working lifetime.
Sustained Attention on One Task	Inspectors may have some tasks where just one defect type is the target, but these are often interspersed with other tasks (e.g. different components) where different defects, often less rare defects, are the target.
Time Uncertainty	Defect occurrence is rarely predictable although inspectors often return to the same area of the same aircraft or engine and attempt to predict when defects are likely.
Spatial Uncertainty	While the actual occurrence of defects at specific places on specific components may be unpredictable, the inspector can have much useful information to guide the inspection process. Training, service bulletins and shared experiences can help point inspectors to specific locations where defects are more likely.
Low Feedback	Aircraft inspectors do not get good feedback, mainly because there is no easy way to find what truly is a signal, especially a missed signal. Feedback on missed defects only comes when one is found at a subsequent inspection, or when an operational incident occurs. Even feedback on false alarms is sporadic. Feedback of both Misses and False Alarms is at best severely delayed and therefore of little use to the inspector.
Unrealistic Expectations	For more common defects, expectations from training can translate relatively faithfully into practice. However, for very rare defects, expectation may still be unrealistically high after considerable practice.
Isolated Inspection Environment	The hangar and even the shop inspection environment are typically noisy, social and distracting. Both noise and social interaction and even some forms of distraction have been found to <u>improve</u> vigilance performance in laboratory tasks.

Table 4. Comparison between attributes of vigilance tasks and aircraft inspection tasks

YEAR 2: DEVELOPMENT OF THE GENERAL AVIATION INSPECTION TRAINING SYSTEMS (GAITS)

Anand K. Gramopadhye, R. Jacob, S. Raina, S. Regunath and R. Subramanian
Advanced Technology Systems Laboratory

Department of Industrial Engineering, Clemson University, Clemson, South Carolina 29364

Abstract: Inspection is an important step in ensuring product quality especially in aircraft industry where safety is the highest priority. Since safety is involved, effective strategies need to be set to improve quality and reliability of aircraft inspection/maintenance and for reducing errors. Humans play a critical role in visual inspection of airframe structures. Major advancements have been made in aircraft inspection, but General Aviation (GA) lags behind. Strategies that lead to improvement in inspection processes with GA environment will ensure reliability of the overall air transportation system. Training is one such strategy where advanced technology can be used for inspection training and reducing errors. A hierarchical task analytic (HTA) approach was used to systematically record and analyze the aircraft inspection/maintenance systems in geographically dispersed GA facilities. Using the task analytic approach a computer based training system (GAITS: General Aviation Inspection Training System) was developed for aircraft inspection that is anticipated to standardize and systematize the inspection process in GA. This report documents the work involved in the development of General Aviation Inspection Training Systems in the GA environment.

INTRODUCTION

Inspection in aircraft maintenance is mostly visual in nature and comprises of 90% of all inspection. Due to this fact the importance of effective human inspection is critical for airworthiness of General Aviation aircrafts. Added to the fact that the aircraft inspection/maintenance being a complex system with many interrelated human and machine components, the significance of ensuring inspector reliability becomes the essence of maintaining an effective and efficient system. Studies in the area of aircraft inspection and maintenance reveal the importance of correct inspection techniques and human decision making performance. Completely eliminating errors committed by the inspectors is always a difficult goal but efforts should be taken to understand the causal factors which lead to error occurrences and emphasis should be laid on training to eliminate the possibility of error occurrence. This report focus on development of a Computer Based Training tool entitled General Aviation Inspection Training Systems (GAITS) designed to help improve the human inspection and decision making performance for aircraft inspection tasks.

TASK ANALYSIS

The development of the GAITS Program followed the classic training program development methodology. As a first step the requirements, needs and goals of the training program were analyzed. Next, a detailed task analysis of the operations was conducted to determine the knowledge, skills, and abilities necessary for the job in order to specify the behavioral objectives of the training program. The

team partners at geographically dispersed GA maintenance sites located within the continental US provided the research team with access to their facilities, personnel, and documentation and allowed the research team to analyze their existing inspection protocol at different times of the shift. The objective of this task analysis was to identify human-machine system mismatches that could lead to errors through shadowing, observing, and interviewing techniques. The goal of the task analysis, which was to understand how the existing system works, was achieved using a formal hierarchical task analytic approach. Table 1 shows a representative task analysis for the search function.

ERROR TAXONOMY DEVELOPMENT

For all inspection functions, the lists of all possible errors were listed and this was mapped using Rouse and Rouse's (1983) error taxonomy to identify the error genotypes. Having this information, expert human factors knowledge was applied to the sub-task to identify specific interventions (e.g., provide job-aids) to minimize the negative effects due to specific error shaping factors and to improve performance on the sub-task. Training needs were developed for producing the correct outcome. As shown in Table 2, the inspection function is "Inspect the frames and structures for cracks, corrosion, loose and for missing rivets". Errors for that particular inspection function were classified using Rouse and Rouse's error classification scheme and training content was established to prevent the occurrence of errors.

YEAR 2 ACTIVITIES:

In year 2, the research team outlined the methods, content and delivery system for use in GAITS. These are described in the paragraph below.

TRAINING METHODS FOR INSPECTION

The basic principles which have been effectively incorporated within GAITS include pre-training, feedback, active training, progressive parts proposes that training should be imparted in a top-down manner, with the general level being taught before the specifics.

Feedback

Accurate and rapid feedback should be provided to the trainees so that they know whether the defects were classified correctly or the search strategies effective. Feedback can be classified as either performance or process. Performance feedback typically consists of information on search times, search errors and decision errors while process feedback provides information to the trainee about the search process, such as areas missed. It has been found that performance can be improved if trainees are provided feedback in the form of knowledge of results coupled with some attempt at performing the task. This is applicable to learning facts, concepts, and procedures as well as to problem solving, cognitive strategies and motor skills. Immediate feedback should be provided at the beginning of the training program, and it should be delayed until the “operational level” is reached. Providing regular feedback beyond the training session helps to keep an inspector calibrated.

Active Training

A trainee should respond actively after each new piece of material is presented by, for example, identifying a fault type or making decision on the degree of a defect. Czaja and Drury (1981) demonstrated the effectiveness of this approach for a complex inspection task.

Progressive Parts Training

Progressive parts training methodology was successfully applied to industrial skills by Salvendy and Seymour (1973). In this methodology, parts of the task are taught to criterion, with successively larger sequences of parts being introduced.

Schema Training

The aim of schema training is that trainees must be able to generalize their training to new experiences and situations. For example, schemas need to be generated projecting every site and extent of the defects found on a plane wing so that the inspector is able to detect and classify a defect wherever it

occurs. Thus, the inspector needs to develop a schema for defects to allow for a correct response in novel situations. The key to the development of a schema is to expose the trainee to controlled variability during training.

Feedforward Training

Feedforward training cues the trainee as to what should be perceived. When novice inspectors try to find defects on an airframe, the indications may not be obvious, unless they know what to look for and where to search.

STRUCTURE AND CONTENT OF GAITS

System specifications and structure

GAITS was developed using Macromedia Authorware 6.5, Macromedia Flash MX and Microsoft Access. The development work was carried on a Pentium(R) 4, 2.4 GHz platform with a 17” resolution monitor, 256 MB RAM, 1.5 MB video RAM, 57.2 GB hard drive and a multi-speed CD drive. The development methodology utilized an integrated task analytic and iterative software development methodology. The training program uses text, graphics, animation, video and audio. The inputs the system are entered through a keyboard and a two-button mouse. GAITS consists of four main modules namely 1) Introduction 2) Training 3) Simulator and 4) Design and Analysis (Figure 1). The software combines graphical user interface technologies along with good usability features. System users interact with the software through a user-friendly interface. Considering ease of use and information utilization, the tool uses a multi-media presentational approach.

Introduction

The Introduction module provides information to the trainee about various facets of the program. It consists of the following

a. Overview: The module gives an overview of the CBT tool. It introduces the trainee to different aspects in the software such as training of search and decision making.

b. Types of inspection: It provides the information about various kinds of inspections, which take place in the General Aviation (GA) environment. In addition to this, different levels of visual inspection are discussed in this module.

c. FAR’s (Federal Aviation Regulations): The module also discusses the FAR’s as they relate to general aviation procedures and guidelines. In addition to this, the introduction module describes the

common tools, which are used in visual inspection, and the factors namely process, physical, subject and organizational, which affect the inspection performance.

Training

The Training module is divided into six units namely Initiate, Access, Search, Decision, Respond and Return (Figure 2), which look into various aspects of the inspection process. The different units, which comprise the Training module, help the trainee understand the conditions, which lead to error occurrence. The module also prescribes correct inspection procedures and steps to prevent error occurrence. Additionally each unit contains a quiz, which checks the trainee's knowledge and the extent to which the trainee has understood the material. (Figure 3 and Figure 4 show the screen shots from the Decision unit.)

Simulator

In order to check trainee's knowledge the simulator provides an utility of simulating an actual structural inspection task. This provides hands on experience in conducting inspection. Additional utility included in the simulator is to check the trainee's performance on the simulated inspection task. The performance of a trainee is tracked using the Design and Analysis.

Design and Analysis

Design and Analysis module enables the instructor to create scenarios to tailor training based on training needs. Moreover it allows analysis of performance scores of the trainee. Once the trainees undergo training in the training module, they can perform actual inspection tasks using the computer simulator. Using the Design and Analysis module the instructor can 1) analyze the results of the students' performance in the training and simulator modules; 2) customize training for each individual. Figure 5 shows how an instructor can create scenarios for wing inspection by selecting alternate images. Based on the performance of the trainee, future scenarios can be developed, such that it helps develop specific inspection skills that are lacking.

DISCUSSION AND CONCLUSIONS

It is anticipated that the use of this training program will result in the following:

Standardization

The use of a computer-based inspection training system eliminates the problems arising from using actual airframe structures and the non-standardization in training resulting from the use of different sets of defects by different instructors. The aim is that all the

trainees will be trained to the same set of standards on the same set of defects.

Adaptability

This computer-based training tool can be tailored to accommodate individual differences in inspection abilities. Images of airframe structures containing defects can be created to train inspectors on particular facets of the inspection task.

Convenience

Retraining can be accomplished more conveniently, and trainees can work on the system whenever they have time available. Also, trainees can work individually, eliminating the intimidation created by a classroom environment or by the presence of an instructor.

Record keeping

The utilities of Design and Analysis allow the instructor to monitor and track individual performance easily. The record keeping process is built into and automated on the software. Individual performance can be tracked initially for training and later for retraining.

Reference:

1. Czaja, S.J. and Drury, C.G., 1981, "Training programs for inspection". *Human Factors*, 23 (4), 473-484.
2. Drury, C.G., Prabhu, P.V., and Gramopadhye, A. K., 1990, "Task analysis of aircraft inspection activities: Methods and findings". *Proceedings of the Human Factors Society 31st Annual Meeting*.
3. Drury, C.G. and Gramopadhye, A. K., 1990, "Training for visual search", In *Proceedings of the 3rd FAA Meeting on Human Factors in Aircraft Maintenance and Inspection: Training Issues*.
4. Drury, C. G., 1991, "The maintenance technician in inspection. In *Human Factors in Aviation Maintenance - Phase One Progress Report* ", DOT/FAA/AM-91/16 (pp. 45-103), Washington, DC: Office of Aviation Medicine.
5. Drury, C.G., 1992, "Inspection performance. In *Handbook of Industrial Engineering* ", (pp. 2282-2314), Wiley. New York.
6. Gould, J. D., Boies, S. J., and Ukelson, J., 1997, "How to Design Usable Systems", In M. G. Helander, T. K. Landauer and P.V. Prabhu (eds.) "Handbook of Human-Computer Interaction",

Elsevier Science B. V., Sara Burgerchartstraat 25 P.
O. Box 211, 1000 AE Amsterdam, The Netherlands.

7. Patrick, J., 1992, "Training Research and Practice", Academic Press, New York.

8. Gramopadhye, A.K., Drury, C.G., and Prabhu, P. V., 1997, "Training strategies for visual inspection", Human Factors and Ergonomics in Manufacturing, 7 (3), 171-196.

9. Gramopadhye, A. K., and Thaker, J., 1998, Task Analysis. In W. Karwowski and W.S. Marras (eds.) "The Occupational Ergonomics Hand Book", CRC Press LLC, 2000 Corporate Blvd., N.W., Boca Raton, Florida 3343.

10. Rouse, W. B., and Rouse, S. H., (1983). Analysis and Classification of Human Errors. In IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-13, No. 4.

11. Salvendy, G., and Seymour, D. W., 1973, "Prediction and Development of Industrial Work Performance", John Wiley, Canada.

12. Wiener, E.L., 1975, Individual and group differences in inspection, In: C.G. Drury and J.G. Fox (eds.), "Human Reliability in Quality Control" (pp. 101-122), Taylor and Francis Ltd., London.

Table 1: Task Analysis

TASK DESCRIPTION	ERRORS	OUTCOME	TRAINING NEEDS
<i>3.1 Search by Fixation in Field of View</i>			
3.1.1 Inspect the frames and structures for cracks, corrosion, loose and for missing rivets.	<p>E3.1.1.1 Does not know how to inspect the frames and structures for cracks, corrosion, loose and missing rivets (EC 5).</p> <p>E3.1.1.2 Does not know how to identify the cracks, corrosion, loose and missing rivets (EC 5).</p> <p>E3.1.1.3 Does not bring the correct tools to inspect the frames and structures (EC 6).</p> <p>E3.1.1.4 Does not inspect the frames and structures for cracks, corrosion, loose and missing rivets (EC 6).</p>	Does inspect the frames and structures for cracks, corrosion, loose and for missing rivets.	Are the inspectors trained on detecting the different type of defects like cracks, corrosion, loose and missing rivets?

Table 2: Error Classification

TASK DESCRIPTION	Task Analysis									OBSERVATIONS	CONTENT
	A	S	P	D	M	C	F	O			
3.0 SEARCH FOR INDICATIONS											
<i>3.1 Search by Fixation in Field of View</i>											
3.1.1 Inspect the frames and structures for cracks, corrosion, loose and missing rivets	X		X		X					Systematically inspected one frame and structure at a time for cracks, corrosion, loose and missing rivets.	Consists information on how to inspect the frames and structures for cracks, corrosion, loose and missing rivets. Consists information on all the different types of defects. Consists information on the tools required to inspect the frames and structures.

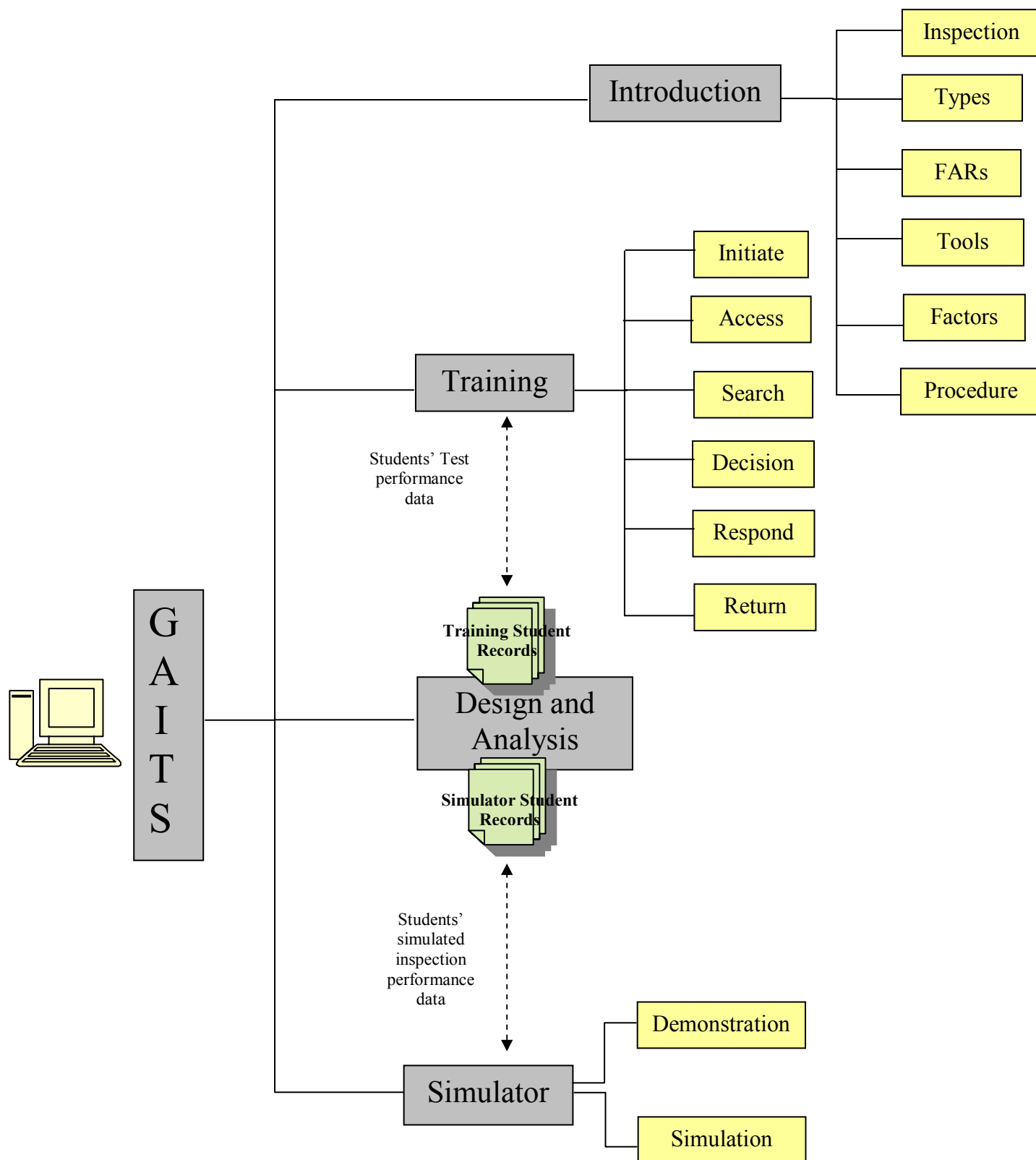


Figure 1: Structure of GAITS

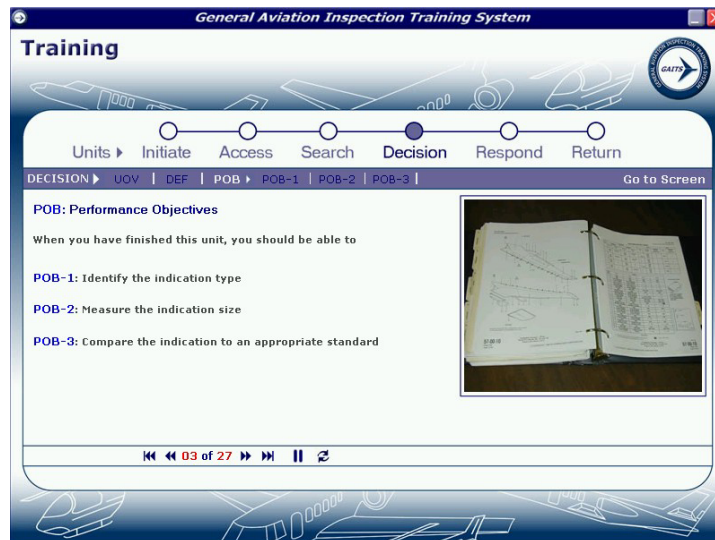




Figure 4: Question slide of the Decision Making unit

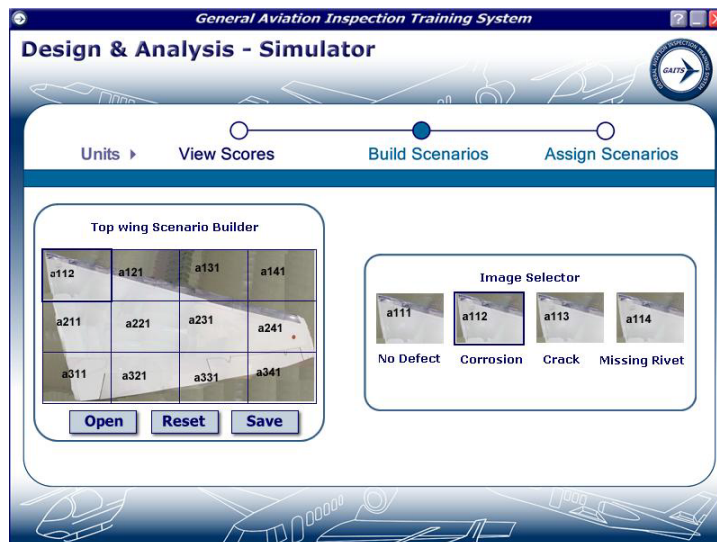


Figure 5: Scenario Builder screen of the Design and Analysis module.

EXPERIMENTS ON LANGUAGE ERRORS IN AVIATION MAINTENANCE: ASIA

Colin G. Drury and Jiao Ma

University at Buffalo, Department of Industrial Engineering
438 Bell Hall, Buffalo, NY 14260
drury@buffalo.edu

The Federal Aviation Administration has raised many issues concerning the outsourcing of maintenance to foreign repair stations and recommends establishing a method for determining whether language barriers result in maintenance deficiencies. This work addresses concerns that non-native English speakers may be prone to an increased error rate that could potentially affect airworthiness. This paper presents Year 2 of the project. We used the seven scenarios of language error developed in Year 1 as the basis for our data collection effort to quantify the frequency of error. An intervention experiment has been designed and tested using a sample of 200 maintenance personnel from countries in Asia. The interventions were found to increase document comprehension performance. Participants tended to maintain a constant accuracy level, with performance changes coming mainly from speed differences.

INTRODUCTION

In 2001, the Federal Aviation Administration raised many issues concerning the outsourcing of maintenance to foreign repair stations in considering changes to domestic and foreign Federal Air Regulations, recommending that:

“The FAA should establish a method for determining whether language barriers result in maintenance deficiencies.”

This project is a direct response to these concerns that non-native English speakers, in repair stations in the USA and abroad, may be prone to an increased error rate that could potentially affect airworthiness. The documentation for repair provided by an English speaking airline is always in English, and this documentation must be used to govern all maintenance tasks, despite a potentially large proportion of mechanics who do not use English as a native language. This paper follows our 2003 HFES paper (Drury and Ma, 2003) and describes the first experiment using a methodology for quantifying the effectiveness of possible countermeasures to language errors.

As noted in our 2003 paper, this project developed seven scenarios of language error based on visits to sites in the USA and the UK; it also provided a model for these unique communication errors based on the communications literature and an analysis of several databases (e.g., NASA/ASRS). Many references to communication theories and studies of outsourcing were given in Drury and Ma (2003) and will not be repeated here.

The seven scenarios found were:

Scenario 1: “The Mechanic (Aircraft Maintenance Technician, AMT) or Inspector was not able to communicate verbally to the level required for adequate performance.”

Scenario 2: “The Mechanic (AMT) or Inspector and the person to whom they were speaking did not realize that the other had limited English ability.”

Scenario 3: “Native English speakers with different regional accents did not understand each others’ communications.”

Scenario 4: “The Mechanic (AMT) or Inspector did not understand a safety announcement over the Public Address (PA) system.”

Scenario 5: “The Mechanic (AMT) or Inspector did not fully understand a safety placard.”

Scenario 6: “The Mechanic (AMT) or Inspector did not fully understand documentation in English, for example a Work Card or a Manual.”

Scenario 7: “The Mechanic (AMT) or Inspector did not fully understand a document translated from another language into their native language.”

In our work, we have been visiting sites worldwide to measure the frequency of these scenarios, and evaluating the effectiveness of countermeasures.

A survey conducted by a major manufacturer showed that language skill varied (as expected) by world region, and that not all sites with lower language skills translated documents into the native language. Our analysis of the survey data reported earlier found that two strategies used to reduce the potential for language errors were (a) translation into the native language, and (b) conducting face-to-face meetings in the native language. However, only about 17% of airlines in the region that most often used translation (Asia) actually translated maintenance documents into the native languages. Even among the group of 8 airlines who reported the lowest English speaking ability, only 2 modified the English documents in any way. Other strategies of intervention found in our site visits included

Thus, we planned to compare four potential language error reduction interventions:

- Some of these methods can be combined, for example the provision of both a Glossary and a bilingual coach, or the addition of AECMA Simplified English to all conditions except for translation into the native language. Finally, for comparison, a baseline condition, no intervention, was required. This paper describes briefly the first two experiments conducted within this framework, and the main data collection in one region, Asia.

Measures

Task Cards

percentage passive voice, and the Flesch-Kincaid reading score. The cards differed on all measures.

Pre-Test Design

Design

Task card Complexity: 2 levels - Simple

- Complex

Task card Language: 2 levels - Simplified English

- Not Simplified English

Language Interaction: 4 levels - No intervention (English)

- English with glossary

- English with coach

- Full Chinese translation

There are several reasons to collect data from MROs located in Asia, especially China. First, in our analysis of the manufacturer's survey data, we found that about 30% of users in Asia had a very limited English speaking ability, another 40% were able to conduct simple conversations; about 40% of the users were able to work effectively with only written maintenance/inspection related documents, and another 15% had very little English reading ability. Compared with North America and Europe, Asia has a much smaller base of English-using mechanics. Second, the Asia-Pacific region is poised to be one of the strongest growth engines for the foreseeable future for the maintenance, repair and overhaul industry (*Overhaul & Maintenance*, 2002). U.S. and European airlines continue to ship wide-body aircraft to East Asia to take advantage of low labor costs. Almost half of the top 10 Asian MROs are located in China. According to *Aviation Week & Space Technology*, "the Civil Aviation Administration of China (CAAC) is confident that despite the downturn in the global

airline industry, more maintenance, repair and overhaul (MRO) joint venture companies will be set up with Chinese airlines within the next two years” (Dennis, 2002).

Participants were tested individually or in small groups. After obtaining Informed Consent and completing demographic questions, the participants were given one of the four task cards and its associated comprehension questions. They were timed, but instructions emphasized accuracy. After the completion of the comprehension task, the participants were given the Accuracy Level Test (Carver, 1987), for the required 10 minutes to act as a potential covariate in our analysis. This test used a total of 100 words with a forced synonym choice among three alternatives, and produced on the scale of reading grade level. It has been validated against more detailed measures of reading level (Chervak, Drury, Ouellette, 1996).

Preparation of the Data Collection Packet for Asia

The translation process took place in two steps. A native Chinese research assistant (9 years as an engineering major), who is very familiar with the task cards, took a lead in translating the packet. A large number of technical and language references were consulted. The principal investigator and other domain experts (e.g., native Chinese mechanical engineers in the Department of Aerospace and Mechanical Engineering at the University at Buffalo, SUNY) were consulted on the technical details (e.g., lockwire). Then both translated task cards, and original packets of data collection material were submitted to a retired professor from the Department of Avionics, Civil Aviation University of China (CAUC) for review.

We developed an English/Chinese glossary for each task card. We had two native English speaking engineering graduate students and two native Chinese speaking engineering graduate students read through all the task cards and circle all the words/phrases/sentences they did not comprehend, or even those about which they were slightly unsure. We built up this glossary to be as comprehensive as possible, including nouns, verbs, adjectives, abbreviations, etc.

Results from Asia: Intervention Performance

This test used 200 participants from six sites in mainland China and Hong Kong. First, in contrast to the pre-tests, there was almost no negative correlation between accuracy and time for the comprehension test ($r = -0.210$, $p = 0.09$). There were moderate correlations of both with Years as an AMT ($p = 0.061$ and 0.008) and Years Learning English ($p = 0.005$ and 0.006). A third measure was created by dividing Accuracy by Time to give a combined overall Performance score.

Reading Level was tested as a covariate, but was not significant in any of three GLM ANOVAs of Accuracy, Time, and Performance. Years as an AMT was a significant covariate in all three measures, but did not change the significance pattern of the three factors, so results of ANOVAs rather than ANCOVAs will be presented here. The surprising overall result was that Accuracy of comprehension did not vary with any of the factors except Site which was included as a main effect only ($F(5, 179) = 2.58$, $p < 0.028$).

The sites were different on Time and Performance measures ($F(5,177) = 7.88$, $p < 0.001$, and $F(5, 177) = 5.46$, $p < 0.001$), with the two sites in Hong Kong being more rapid

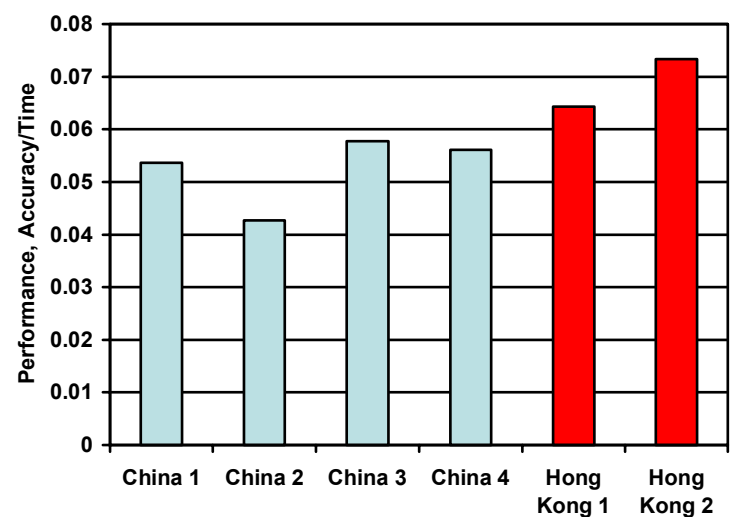


Figure 1: Performance Results by Site

and having a higher performance than the mainland China sites (Figure 1).

The other significant main effects were for Intervention and Task Card. Intervention was significant for Time ($F(3,179) = 7.57$, $p < 0.001$) and Performance ($F(3,179) = 2.99$), while Task Card was significant at ($F(1,179) = 15.43$, $p < 0.001$) and ($F(1,179) = 5.02$, $p = 0.026$) respectively. The Easy task card had a performance score of 0.058 while the Difficult task card scored 0.052. Interventions grouped into two sets, with all three active interventions faster than the baseline condition. In terms of Performance, the comparisons are shown in Figure 2. Note that the use of AECMA Simplified English had no significant effect on any measures. Also, no interactions among any factors reached significance, simplifying the interpretation of results.

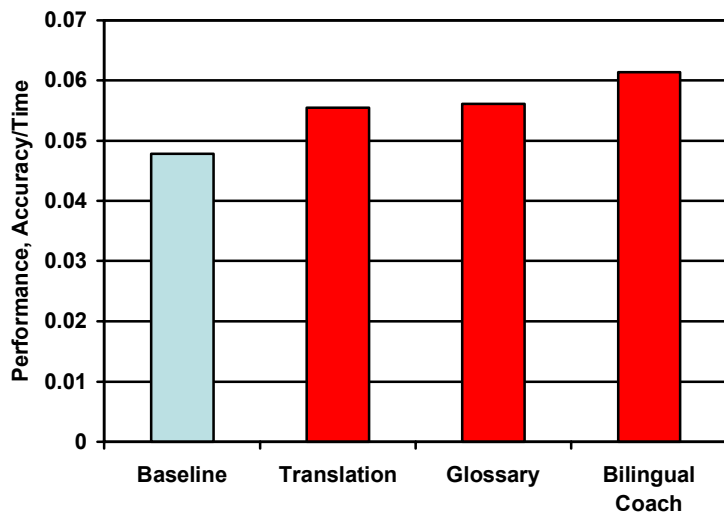


Figure 2: Performance comparisons by Intervention.

Results from Asia: Scenario Analysis

In addition to the evaluation of the interventions, we used a questionnaire to determine the relative incidence of the seven scenarios developed earlier. A number of measures of incidence were used, including estimates of the time since last occurrence. Here we present only the overall response to “Have you ever encountered an error of this type?” The incidence of each scenario is shown in Figure 3 for mainland China and Hong Kong separately. Note that the four most frequently encountered scenarios (1, 2, 6 and 7) are concerned with directly work related verbal and written ability. The other three scenarios concern regional accents, and less-work-related events. The misunderstood translations (Scenario 7) often referred to translations from English by aircraft manufacturers for whom English is not the native language.

For the response to factors most associated with these scenarios, GLM ANOVA of the percentage encountering each incident by Factor was performed, with Region and Scenario as additional independent variables. All main effects and interactions except Factor × Country were significant at $p < 0.01$ or better. The responses divided into two groups, one seen as highly related to the incident and one less related. These are:

Highest Related to Scenarios

- Task is Complex
- Task Instruction is complex
- AMT’s inadequate written English
- AMT’s inadequate verbal English
- Time pressure on AMT

Lowest Related to Scenarios

- Poor communication equipment
- AMT does not ask for help
- AMT uses native language under stress
- Unwilling to expose lack of English

A similar analysis was performed for the ten factors potentially mitigating language errors. The GLM ANOVA gave significance at $p < 0.01$ for Factor, Region, and their interaction. As with causal factors, the results grouped into two:

Highest Related to Scenarios

- Translated documents
- Consistent terminology
- Document uses good design practices
- Use of aircraft for communication
- AMT is familiar with the job

Lowest Related to Scenarios

- AMT has passed comprehension test
- AMT is certified for that job
- Translator is available to AMT
- Jobs is assigned based on English ability
- AMT team with English speaker

As with causal factors, the highest group included the physical changes, plus in this case job familiarity. The lowest group was mainly individual and social interventions.

Finally, an analysis of how errors are discovered was performed. Only Scenario, Factor, and the Factor × Country were significant (at $p < 0.02$). Again, there was a grouping of the Factors, this time into 3 groups:

Highest Related to Scenarios

- AMT asked for assistance/clarification

Medium Related to Scenarios

- AMT appeared perplexed
- Resulting physical error was detected

Lowest Related to Scenarios

- AMT agreed with everything said
- AMT did not understand at buy-back
- AMT closed access prematurely

From these groupings, note that the least commonly found were either an unusual behavior, or events later in the maintenance/inspection process.

CONCLUSIONS

On our site visits, we conducted two main studies. The first was a direct test of the effectiveness of four interventions and the second an evaluation of the incidence and causal factors in seven previously-developed language errors scenarios.

The interventions experiment used a baseline condition of English documents, and then added translation

(including the test form), a glossary, a bilingual coach, and a combination of these last two conditions. We used two levels of task card difficulty, each with and without Simplified English. This made a three-factor factorial experiment (Intervention \times Difficulty \times Simplified English), with various covariates.

On the samples tested so far, the results are encouraging. While there were some differences between regions, differences between interventions were consistent across regions. All of the interventions had some effect, although mainly on the times and our performance measure, rather than on accuracy *per se*. If this indeed reflects practice, then maintenance personnel appear to slow down when they find language difficult, rather than making more errors at a constant speed.

The analysis of the incidence and factors data suggest that most of the causal factors in language errors are seen to be either directly document-related or time pressures. The factors least related are much more behavioral or communications channel related. A similar result was found for mitigating factors. These findings give some credence to our use of the documentation interventions, which should address four of the five highest related causal factors.

Our next task is to repeat this experiment in other continents. The current plan is to visit locations in Central and South America in Fall 2004 and Europe in Spring 2005.

REFERENCES

Carver, R. P. (1987). *Technical Manual for the Accuracy Level Test*. REV-TAC Publications, Inc.

- Chervak, S., and Drury, C. G. (2003). Effects of Job Instruction on Maintenance Task Performance. *Occupational Ergonomics*. Vol.3, Issue 2, 121-131
- Chervak, S., Drury, C. G., and Ouellette, J. L. (1996). Simplified English for Aircraft Workcards. *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*, 303-307.
- Dennis, W. (2002). MRO to Grow in China. *Aviation Week and Space Technology*.
- Drury, C. G. and Ma, J. (2003). Do Language Barriers Result in Aviation Maintenance Errors? *Human Factors and Ergonomics Society 47th Annual Meeting Proceedings*, Denver, Colorado, October 13-17, 2003.
- Drury, C. G. and Ma, J. (2004). Experiments on Language Errors in Aviation Maintenance *Human Factors and Ergonomics Society 47th Annual Meeting Proceedings*, New Orleans LA September 20-24, 2004.
- Drury, C. G. and Ma, J. (2003). *Language Errors in Aviation Maintenance: Year 1 Interim Report*, Reports to William J. Hughes Technical Center, the Federal Aviation Administration under research grant #2002-G-025.
- Drury, C.G., Wenner, C., and Kritkauskys, K. (2000). Information Design Issues in Repair Stations. *Proceedings of Tenth International Symposium on Aviation Psychology*. Columbus, OH.
- MRO in Asia-Pacific Region (2002). Aviation Week's ShowNews Online: www.aviationnow.com. Online resources: <http://www.awgnet.com/shownews/02asia1/mro09.htm> & *Overhaul & Maintenance*.
- Patel, S., Drury, C. G. and Lofgren, J. (1994). Design of Workcards for Aircraft Inspection. *Applied Ergonomics*, 25(5), 283-293.

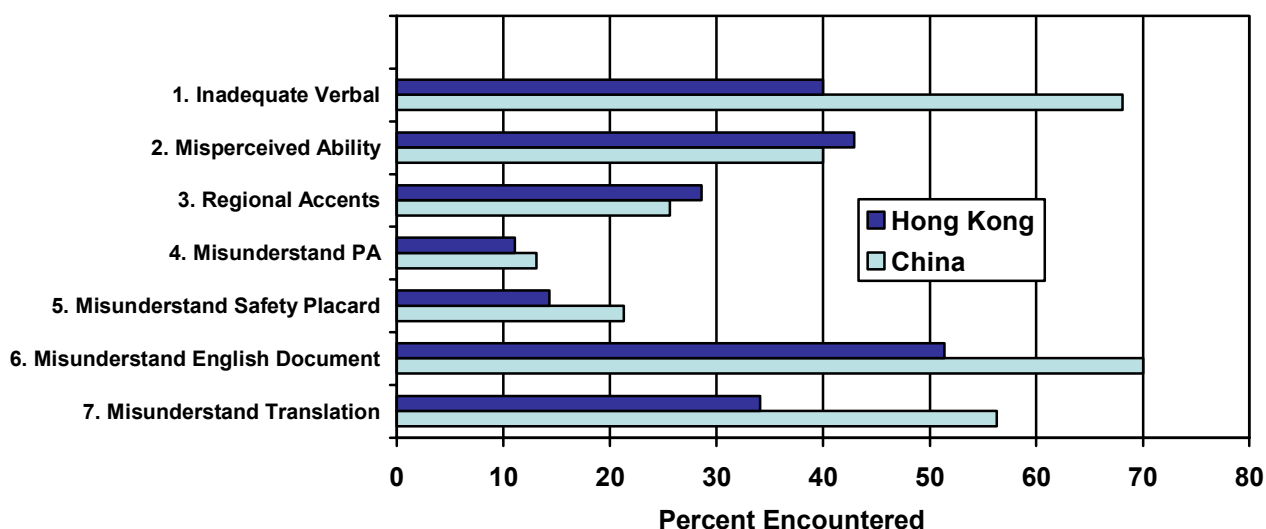


Figure 3: Relative frequency with which each of the seven scenarios was encountered.

EMPIRICAL BASIS FOR AN OCCUPATIONAL VISION STANDARD

Bettina L. Beard¹, Karen M. Jones², Cynthia Chacon² and Albert J. Ahumada¹, Jr.

¹NASA Ames Research Center, Moffett Field, CA 94035-1000,

²San Jose State University Foundation, San Jose, CA

ABSTRACT

There are no government mandated vision standards for aviation maintenance inspectors. Empirically derived vision standards for other occupations cannot be extended to this very different occupation. We apply a psychophysical human-in-the-loop methodology toward defining an empirically-based visual acuity standard for a representative task performed by aircraft maintenance inspectors. Visual acuity declines are simulated using a Gaussian blur function on airframe images. Psychophysical data were collected in non-inspectors and in highly experienced aviation maintenance inspectors. The data may be used to construct an empirically-based visual acuity standard.

INTRODUCTION

It is difficult, if not impossible, to eliminate human error in the process of inspection. Interventions must be developed to reduce these errors and make the process more error-tolerant. Since visual inspection represents a large part of aviation maintenance inspection, one mitigation strategy is to define vision standards for this vision-intensive, safety-critical occupation. A fine-tuned ability to localize, detect, discriminate, and identify job-relevant stimuli can bring cost savings and safety benefits to industry.

In 2001, an FAA Advisory Circular (AC No: 65-31) recommended examination guidelines for the vision of non-destructive inspection (NDI) personnel. It was suggested that near and far vision in at least one eye must be 20/25 and 20/50, respectively. Both near and far requirements could be with corrected or uncorrected vision. This FAA recommendation was based on acuity standards defined in other NDI/NDT occupations. Reviewing the occupational vision standards literature, Beard *et al.* (2002) found no studies that allow generalization of standards to aircraft maintenance inspection. It is unknown how similar tasks must be to validly borrow standards from another occupation without being subject to compromise. What is needed is a rapid, empirically-based methodology for defining occupational vision standards.

No current general standard exists in the aviation industry for the visual qualifications of aircraft maintenance inspectors. Some maintenance facilities use the visual acuity and color vision standards suggested in the FAA Advisory Circular, while other facilities have defined their own vision requirements. This illustrates the need for a uniform and universally accepted set of vision standards that would apply to all aircraft non-destructive inspection and testing (NDI/NDT) personnel.

There are several broad steps that should be taken toward setting an objective, empirically-based occupational vision requirement. The first step is a thorough vision task analysis. In the current context, the FAA commissioned CAMI to perform this analysis focusing on the role of visual processes. Next, to see if a rigorously defined standard can be borrowed from a similar occupation, a review of the literature should be undertaken. Beard *et al.* (2002) compiled a review of a text and WEB-based search for occupational vision requirements, knowledge gained from site visits to major aircraft maintenance facilities, relevant information from technical, mechanical, and inspection textbooks, the FAA maintenance human factors web-site¹, and the human vision literature.

If the standard cannot be legitimately borrowed from a previous standard, an objective research methodology should be followed. In their review of the vision standards literature, Beard *et al.* (2002) identified four occupations that had empirically derived standards. These empirical methodologies ranged from mathematically measuring the size and working distance of the critical visual details (Sheedy, 1980) to psychophysical measurements with blurring lenses placed in front of the eye on a single task (Good & Augsburger, 1987; Good *et al.*, 1996) or multiple tasks (Padgett, 1989).

Here we present a strategy for defining a visual acuity standard that permits increased experimental control by blurring the image before presenting it to the observer, within a computer program. In this way what is done to the signal is exactly known. On the authors WEB page

¹ <http://hfskyway.faa.gov>

(<http://vision.arc.nasa.gov/personnel/tina/beard.html>) we provide the software so that this methodology may be used toward setting standards in other visually intensive occupations.

The primary objective of this research is to aid in the development of recommendations for visual acuity requirements for aviation inspection personnel. Specifically we determine that visual acuity deficits reduce critical task performance and show in graphical form the relationship between acuity decline and performance.

METHODS

Choice of a critical vision task for inspectors

The central question that must be addressed is “At what level of visual deficit would a maintenance or inspection worker become unable to safely and efficiently perform the critical visual tasks required by the job?” Aircraft inspection is a complex process, requiring many tasks, skills, and procedures. There are multiple critical vision tasks that the workers are required to perform. One purpose of inspection is to detect surface discontinuities such as cracks within the airframe and powerplant regions of the aircraft. Cracks are typically caused by two surfaces being overlaid at a boundary (Hellier, 2001). Since these cracks may be very small and of low contrast, adequate visual acuity is likely to be involved in their detection. After consultation with domain experts, crack detection was chosen as the representative task in order to ultimately set a visual acuity standard for aircraft maintenance inspection.

Psychophysical Experiments

Observers

Two female non-inspection personnel (age range from 23-30) and seven male maintenance inspectors (age range from 35-58 years) participated in the study. Maintenance inspectors were actively employed and had from 10-18 years on the job. All wore corrective lenses, though not always while inspecting. Near and far visual acuity, stereo vision, and color vision tests revealed that all had at least 20/20 acuity, good color vision, however one inspector lacked stereo vision. The inspector lacking stereo vision did not differ significantly from the other inspectors in overall detection performance (data shown below).

Stimuli

Airframe and powerplant crack images were obtained from various sources. Color images were converted to 8 bit black-and-white images to delete any color cues. Before the experiment, “crack removed”

stimuli were generated. Using Photoshop™, the crack was deleted from the image while maintaining the integrity of the background image. The 15 images used in the current experiment are provided in Appendix A.

A “background-with-crack” image at a particular contrast level was generated by multiplying the full contrast difference image (the crack itself) by a multiplicative factor (≤ 1) and adding it back to the background image. The contrast in dB is 20 times the log to the base 10 of the factor. An image with the contrast of 0 dB has the original crack. An image with a crack contrast of 6 dB has the difference image reduced by a factor of 0.5. This logarithmic scale keeps the variation in the results more constant over different threshold levels.

Crack length estimation

To accurately determine the crack length and width, estimates of the magnification in each photo had to be determined. Each photo included a circular label or ‘sticky’ whose diameter is a known 0.75 in. To estimate the image magnification, Photoshop™ was used to identify the coordinates of six points along the perimeter of the sticky. These estimates of the perimeter were taken by eye; therefore the error in these judgments was also determined. A computer program took these data and computed a magnification value estimating the diameter of the sticky. When the sticky was on a flat surface, the image is an ellipse and the estimates were very accurate. Some of the stickies were located on an edge or curved surface. In these cases, coordinates were identified only on the flat portion of the sticky and the ellipse estimated based on this flat portion.

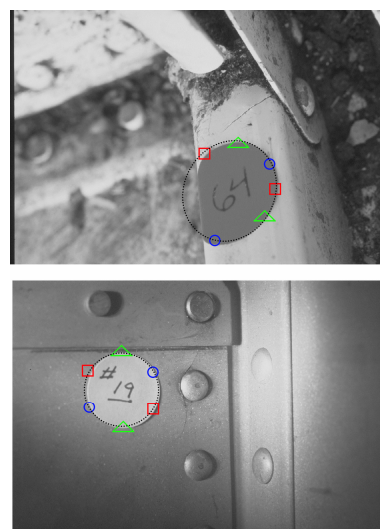


Figure 1: Crack length and width estimates. Each photo included a circular label or ‘sticky’ whose diameter is a known 0.75 in. A magnification value

estimating the diameter of the sticky were computed from six points along the perimeter of the sticky (shown in the figure).

At first, images were adjusted so that all the stickies had a diameter of 0.75-in on the experimental display screen (the same as the sticky's actual size), but these images were so coarse because of display resolution limitations that features of the fine cracks disappeared. Images were then adjusted to a screen sticky size of 3-in, resulting on average image width reduction from 1500 pixels to 800 pixels. Some of the images were still larger than the screen resolution of 1024 by 768 and so were cropped to 990 x 660.

Apparatus

Photographs of large engine airframe cracks were presented on a 1024x758-pixel display screen (SONY Trinitron). Viewing was binocular with natural pupils. From observations of aircraft inspectors performing primary inspections, Goode (personal communication) found that the majority of visual observations were done in the distance range from 34 to 40 cm. Because of screen resolution limitations, images were magnified by 4 as discussed above and so the experimental distance was comparably increased to 160 cm. From this distance each pixel subtended 0.31 arc min. The display background screen had a mean luminance of approximately 40 cd/m². Three lights illuminated a gray wall behind the monitor. Another lamp illuminated the ceiling behind the observer to achieve ambient lighting. Photometric measurements of the SONY monitor revealed that screen luminance values remained constant only after it was turned on for at least 45 minutes.

Simulating Visual Acuity Decline

Although the shape of the human blur function differs between individuals and changes for different optical conditions, it can be approximated by a Gaussian blur function. An observer with 20/20 visual acuity was assumed to have a Gaussian blur spread² of 2 arc min (Barten, 1999; Ahumada, 1996). A person is said to have 20/40 visual acuity if they see at 20 ft what a 20/20 person sees at 40 feet. If we assume that the 20/40 person has the same contrast sensitivity as the 20/20 person, then the blur for the 20/40 person must be twice the blur of the 20/20 person. Therefore, to simulate 20/40 visual acuity the combined blur of the image and the observer should be 4 arc min. The combination rule for

Gaussian blur is the Pythagorean rule, so, for example, to obtain an acuity value of 20/40, the image blur spread was set to 3.46 since the $\sqrt{3.46^2 + 2^2}$ is 4. To obtain an acuity of 20/A where A = the desired acuity level, then the blur in minutes = $2 \sqrt{((A/20)^2 - 1)}$. Figure 2 presents example "crack removed" and background-with-crack images with and without blur.

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Figure 2: Examples of the crack-removed (upper left panel) and background with crack (lower left panel) images. The two right panels demonstrate these images after they have been blurred to simulate visual acuity decline.

Procedures

Crack Contrast Detection Thresholds

To increase the number of images tested and the range of conditions, the two non-inspector observers collected data on a large set of crack images at a greater number of blur levels, while the NDI/NDT inspectors were run on subsets of crack images and blur levels.

Contrast detection thresholds were obtained using a two interval forced choice staircase method. The background airframe image remained on during the duration of the block of trials. On a single trial, observers were presented with the background alone in one 500 msec time interval and the background with crack in another 500 msec time interval. The interval containing the crack was randomized. The two time intervals were demarcated with a simultaneous tone. Interval one contained one tone burst, while interval two contained two tone bursts. Only one of the time intervals contained the crack stimulus. The observer's task was to choose which interval contained the crack stimulus. The inter-stimulus interval was 500 msec. The sequence of each block of trials and the crack with background image were randomly chosen.

A different airframe image was presented in each

² The spread is the distance from the center to where the blur amplitude is 1/e (0.3679) of the center amplitude.

block of trials, selected by a random permutation of all of the images and blur levels to be presented in a replication, for at least three replications. To help the observer find the crack, in initial practice trials the crack position was indicated to the observer by surrounding the crack with a rectangle. After localizing the crack, the observer could then practice the crack detection task without the surrounding rectangle before continuing on to the experiment.

On the first trial of a block of trials, the crack stimulus was presented above threshold. Estimates of these supra-threshold contrast levels were determined from model predictions (see Ahumada & Beard, 1998) and pilot data. The contrast was adjusted by a staircase procedure. On each trial, if the observer correctly responded as to which interval the crack was shown, then the response was tallied as correct. After three consecutive correct responses, the crack contrast was decreased by a specified amount (step factor). If the observer chose the interval that did not contain the crack stimulus, then a brief feedback tone would sound, the response was tallied as incorrect, and the crack contrast increased by a specified amount on the next trial. To more rapidly converge to threshold, initially the contrast step factor was 2 dB, but was reduced to 1 dB after a change in the direction of the staircase (a reversal), and then reduced to 0.5 dB after the second reversal. After eight reversals in contrast and at least 30 trials, but no more than 50 trials, the block of trials was terminated and the detection threshold calculated by a probit analysis for that crack with background image.

The two non-inspectors collected data on 10 images. The seven highly experienced aircraft maintenance inspectors collected data on either a subset of these same 10 images or on 5 different images. Observer CA collected data on images that represented six levels of blur or acuity levels: 20/20, 20/25, 20/30, 20/35, 20/40, and 20/50. Observer KJ ran on this same set of acuities plus an acuity level of 20/45. The 7 maintenance inspectors collected data on 4 acuity levels: 20/20, 20/30, 20/40, and 20/50. To

evaluate the effect of viewing distance on the detection thresholds, one NASA observer was run on a subset of her conditions at a farther viewing distance of 267 cm.

Contrast Sensitivity Functions

To estimate the observer's internal blur and screen resolution limitations, each observer's contrast detection thresholds were measured for a range of stimuli. The Contrast Sensitivity Function (CSF) provides an estimate of visual acuity because an individual's resolving power is indicated by the intersection of the curve on the abscissa of the graph. Horizontal and vertical thresholds were obtained to estimate meridional differences in the amount of blur within the experimental display.

Much like the experimental task, the observer had to decide in which of two 500 msec intervals the stimulus was presented and respond accordingly (i.e., they responded by pressing '1' if they thought the stimulus was presented in interval one, and '2' if they thought the stimulus was presented in interval two.) There was a 300 msec gap in between the presentation of the two stimulus images. Instead of cracks, however, the target stimuli for this experiment were a square, line and a dipole. Observers completed this experiment while sitting 273 cm from the screen.

RESULTS

Probit analyses were done on each block of trials to estimate the contrast threshold, the value at which the probability of correctly identifying the interval was 75%. The median of the scores replicating a particular condition was then computed.

In Figure 3, detection thresholds are presented across blur or simulated acuity levels. Each symbol represents a different airframe image. The data for each image were fit with linear functions with slopes ranging from -1.3 to -2.9 (median slope = -2.2) for Observer CA and from -2.0 to -3.3 (median slope = -2.9) for Observer KJ.

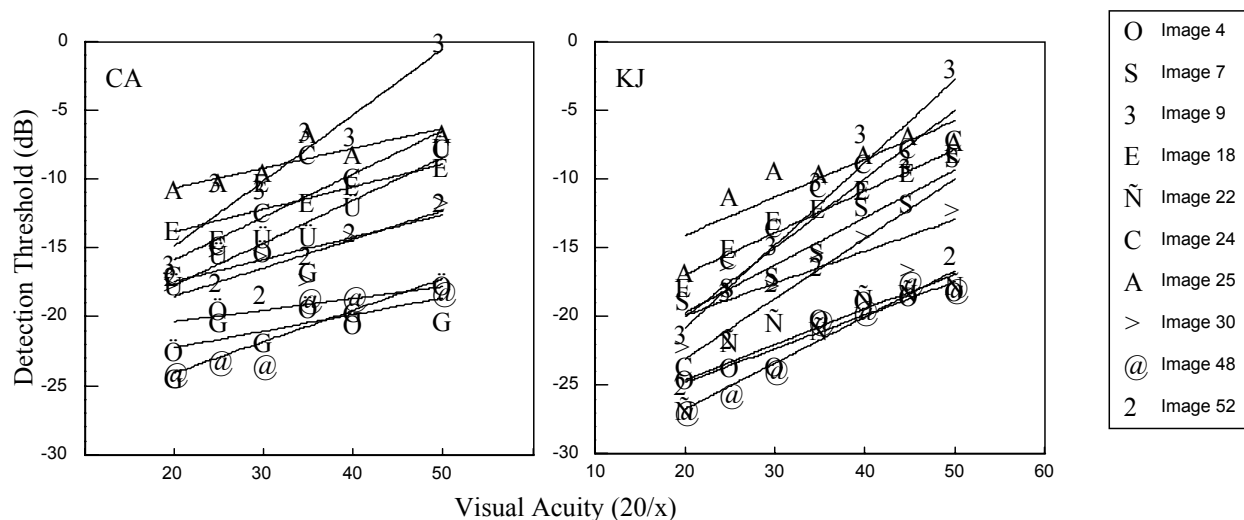


Figure 3: Contrast thresholds are presented across blur or simulated acuity levels. Each symbol represents a different airframe image. The results for observer CA and KJ are shown.

Figure 4 presents contrast thresholds for the different images as a function of blur averaged over the nine observers (inspectors and non-inspectors). Each symbol represents a different airframe image. There is a general tendency for the effect of blur to be larger as the thresholds increase. The two images with the highest thresholds could not even be run at the higher blur levels. Again, the data for each image were fit with linear functions with slopes ranging from -1.5 to -2.8 (median slope = -2.3).

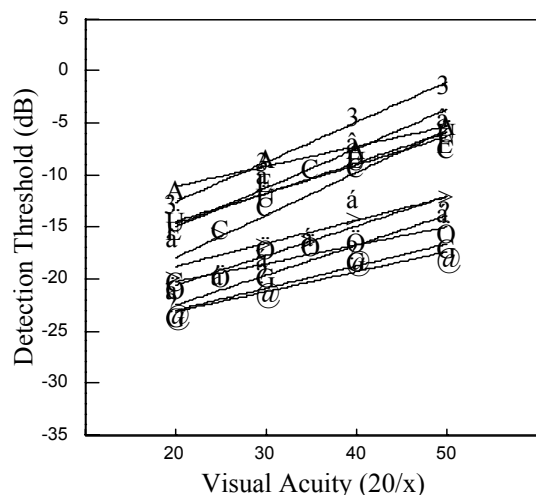


Figure 4: Contrast thresholds for the different images as a function of blur averaged over the nine observers (inspectors and non-inspectors). Each symbol represents a different airframe image (legend is the same as in Figure 3).

Figure 5 shows the effects of blur on observers averaged over images. There is a general tendency for the effect of blur to be greatest for the observers

with the lowest thresholds. The two non-inspectors (CA and KJ) showed lower detection thresholds than did the experienced aircraft inspectors. The reason for this is that observers had participated in a study of practice effects on contrast thresholds in a complex scene (Beard, et al., in preparation) and therefore are highly experienced psychophysical observers.

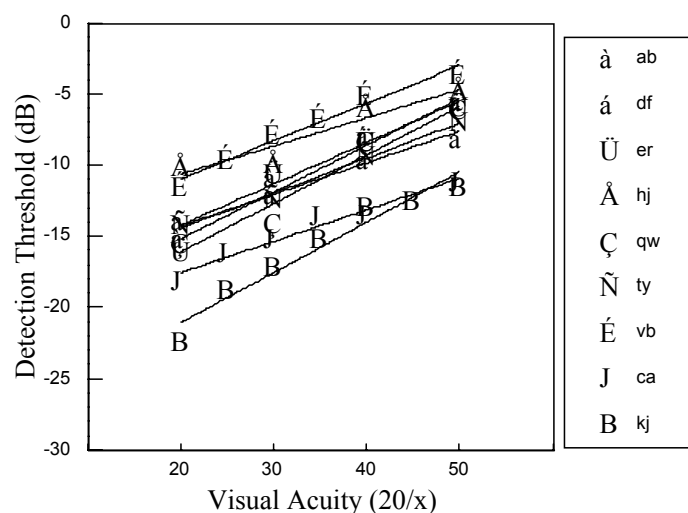


Figure 5: Effects of blur for each observer averaged over images. Observer initials are shown in the legend.

All data presented thus far were collected at a distance of 160 cm. Because not all inspections are done from one single distance, thresholds were measured from a second distance of 266.8 cm. Thresholds were elevated at a further distance, and show a similar increase in threshold with increases in blur.

Figure 6 shows the effect of increasing the viewing distance. Thresholds for the far distance are consistently higher than those for the nearer distance. If the detection were simply a function of target contrast energy, the threshold would be expected to increase by $20 \log_{10}(267/160) = 4.4$ dB. Attenuation of the high spatial frequency energy should cause an additional increase in the threshold, which should be greater for the less blurred stimuli and the higher threshold stimuli.

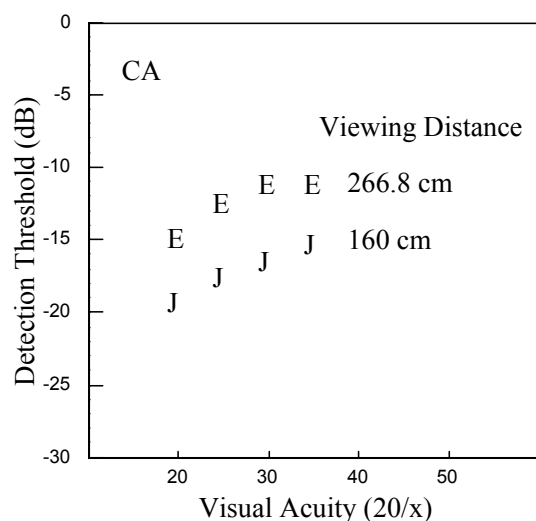


Figure 6: Viewing distance effect. Data were collected in one non-inspector observer.

To foster translation of these data into an occupational visual acuity standard, in Figure 7 we have transformed the data from Figure 5 into Probability of Detection (PoD) curves. The data were converted back to probability of Yes/No detection after being normalized by setting the

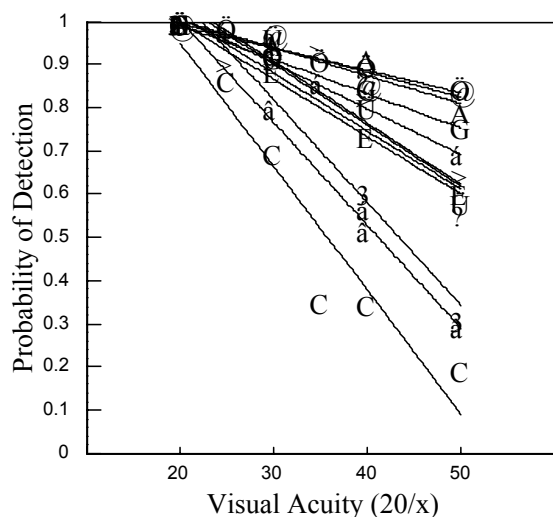


Figure 7: Probability of Detection curves.

unblurred probability of detection to 0.99. This calculation depends strongly on the assumed slope of the psychometric function. Here we assume the standard deviation of the cumulative Gaussian to be 4 dB, but the actual value could be anywhere from 1 dB to 6 dB.

DISCUSSION

Although good vision is a vital qualification for aircraft maintenance inspectors, no general standards for visual acuity currently exist for this occupation. Vision standards from other occupations cannot be “borrowed” to set a standard for maintenance inspectors because the visual demands between occupations are dissimilar and the majority of occupational vision standards are not empirically based (Beard *et al.*, 2002).

One way to look at the effect of not having 20/20 vision is to say the an inspector with 20/40 vision sees at 20 feet what the 20/20 inspector sees at 40 feet. That is to say that the 20/40 inspector has to be twice as close as the 20/20 inspector to make the same discriminations. When the viewing distance is halved, the foveal search area is reduced by a factor of 0.25, so it would take about 4 times as much time to search the same area with the same discriminative ability.

In this project we measured detection performance on a representative task performed by aircraft maintenance inspectors as a function of image blur. These measurements allow predictions of the amount the probability of detection could change as a function of blur. As shown in Figure 7, cracks whose detection was initially at 99% could be greatly reduced by blur corresponding to only 20/30 if the inspection situation was kept constant in all other respects.

The amount of visible contrast energy in the crack correlated well with the contrast thresholds for the crack ($r = -0.89$). However, the effect of the blurring on the thresholds was much greater than the loss in visible contrast energy. For the two images with the greatest loss in visible contrast energy (4.7 dB) at the 20/40 blur level, the average threshold loss was 10 dB. Although this may be in part due to a lack of experience with these blurred images, it is also possible that the blur causes more problems with crack detection than predicted by contrast energy loss alone, such as affecting the extraction of edges. The loss in visible contrast energy can be thought of as a lower limit for the effect of blurring.

Blurring is only one possible cause of lowered acuity. Another possible cause is decreased overall contrast sensitivity. In this case, the predicted effects are

expected to follow more closely the rule that a cutting the viewing distance in half will compensate for a 6 dB loss in sensitivity.

Methodological Limitations and Strengths

The experimental image generation procedure was only an approximation of actual visual inspection. Inspectors were able to use only one very relevant strategy (contrast detection) to look for the defect embedded within a number of realistic aircraft locations. Although the cracks were positioned on actual aircraft structures, inspectors could not use many of the common strategies used in their work environment, such as tribal knowledge (knowing where to look), moving closer, use of shadows (i.e., changing the angle of light from their flashlight), touching the crack. But there is a trade-off between being able to use these techniques and the time it takes to do a search. Differences between the background conditions indicate the effect of background variations on performance and will reduce the importance of decision strategies on defect detection. This methodology permits manipulation of defect absence, length, color, and other attributes. It is important to be capable of manipulating the absence of a defect since uncertainty plays a large role in maintenance inspection (i.e., there is no prior knowledge that a defect will be present). In fact, it is only occasionally that a defect is actually present.

Vision is a fundamental component of effective aircraft inspection. All the same, so too are other cognitive factors such as attention, memory, and experience. Inspectors are knowledgeable about individual components as well as the overall aircraft being inspected, thus they possess the background to properly locate, identify, and evaluate aircraft defects. Often NTSB accident reports will point at visual deficits as contributors to accidents because a crack went undetected, or a worker failed to detect fatigue damage. However, it may not be that vision led to these overlooks. Other cognitive factors may have played major roles in the lack of detection: job-related stress, worker fatigue, multi-tasking, or memory effects of interruptions. The proposed research isolates vision requirements on these duties. Because the job entails much more than vision, these results may not relate to how well the inspector will do on the job. Therefore, although vision is a critical component in inspection, other factors weigh in heavily in the naturalistic task.

Other requirements should address the effects of other cognitive contributors. These data can then be used by the FAA to write acceptable cognitive and perceptual standards and procedures for inspectors including the type and frequency of vision testing

necessary to ensure the safe and effective performance of current employees and job applicants who will perform a particular inspection procedure.

Although psychophysical human-in-the-loop experiments can provide accurate and objective data toward setting a standard, it would be optimal to be able to predict performance using a computational model. Ahumada & Beard (in preparation) show that a model of image discrimination does predict similar blur effects as reported for model predictions of simulated crack stimuli (Beard *et al.*, 2003) but under-predicts the blur effects seen in psychophysical data using these actual crack stimuli.

Guidance toward the setting of a standard

These measurements do not provide a standard, but it converts the problem to specifying a desired physical limitation in performance. The final step in the process of defining a visual acuity standard lies in the hands of the FAA. Using the data in Figure 7, the FAA must decide which stimulus characteristics and what margin of error (e.g., 1 error in one million) will define where to draw the line for the standard.

Recruitment, testing, selection, and training costs are high. The rejection of qualified persons imposes an unnecessary cost on maintenance facilities. While the failure of proper performance on visual tasks could be catastrophic, persons with refractive errors such as correctable myopia who can perform the job should be permitted to do so. Vision requirements should be based on a demonstration that, for example, 20/25 near or 20/50 distance visual acuity is actually needed to perform the essential task. If the task is not generally performed alone (i.e., there are several people in close proximity who provide assistance) then these tasks should not be imposed with a vision requirement for all the individuals. In addition, vision requirements must be based on tasks that cannot be modified by current available technology to assist the vision of the worker.

The governing body, here the FAA, should clearly define the purpose of any vision test and not provide medical examiners considerable latitude when conducting visual acuity testing and evaluation. An interesting case where this was not done, highlights the importance of this recommendation. In a Safety Advisory entitled ‘Determination of Vision Impairment among Locomotive Engineers’ (SA-98-1) published by the Federal Railroad Administration (FRA) and the Department of Transportation (DOT), a lesson can be learned for the current purpose. The FRA’s expectation was that the physicians who would be designated as railroad medical examiners would be trained to competently administer color vision examinations. Thus, they did not anticipate

that it would be necessary to specify for the medical examiners the test procedure to be employed when testing for whether a person meets the standards specified in this rule. That assumption has been called into question under tragic circumstances. If the current rule had been implemented as the FRA expected, the rule would have adequate to prevent a major railway accident involving the fatal collision between two New Jersey transit commuter trains (NTSB/RAR-97/01). The NTSB report found that the medical history of the suspect engineer showed that he had been administered an acceptable test annually by the same contract physician for over 10 years. In the tenth year, the test results showed a deterioration of the engineer's ability to distinguish among some colors. The engineer was then given a Dvorine Nomenclature Test to further evaluate his color vision. Many color weak individuals can identify the names of colors by their brightness instead of their hue. The examiner failed to administer the accompanying Dvorine Second edition color vision test, which measures color discrimination abilities and therefore the results of the first test suggested that the engineer did not have a problem. It was ruled likely that the accident was preventable if the physician had used a sound approach to measure the person's ability to distinguish colors.

Self-monitoring

Aircraft maintenance inspectors as a group take great pride in their ability to detect defects. In addition, they care deeply about the safety implications of their job. Many environmental and developmental variables can affect visual sensitivity. Changes in vision are typically slow and subtle and therefore not easily identified by the individual. Long work shifts or age-related accommodative changes can lead to eye strain, headaches, excessive rubbing of the eyes, esotropia or exotropia, and reduced efficiency on the job. Without an objective measuring tool, workers will not detect gradual changes in their vision. If you don't see something, you don't know that you can't see it (self-awareness). Providing the workers with a method to self-monitor their visual acuity would enhance occupational safety and safety in the NAS.

ACKNOWLEDGMENTS

Supported by FAA Agreement DTFA-2045 to B. L. Beard.

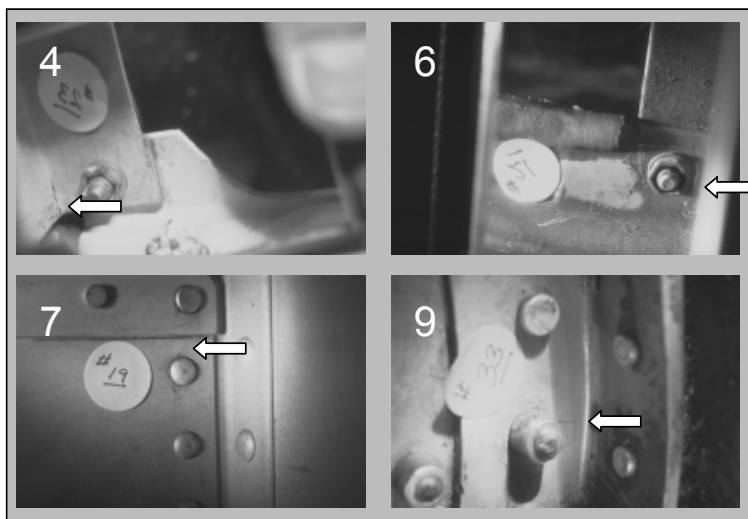
We thank the inspectors who participated in the study. They were wonderful psychophysical observers who always showed great professionalism and devotion toward this effort.

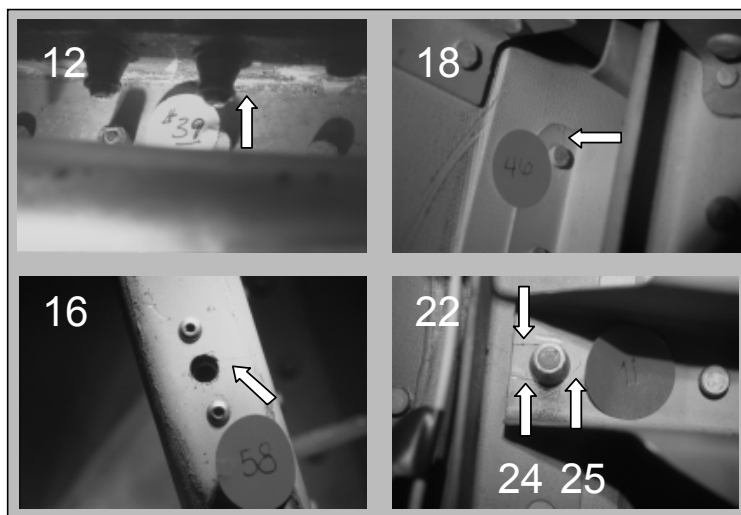
REFERENCES

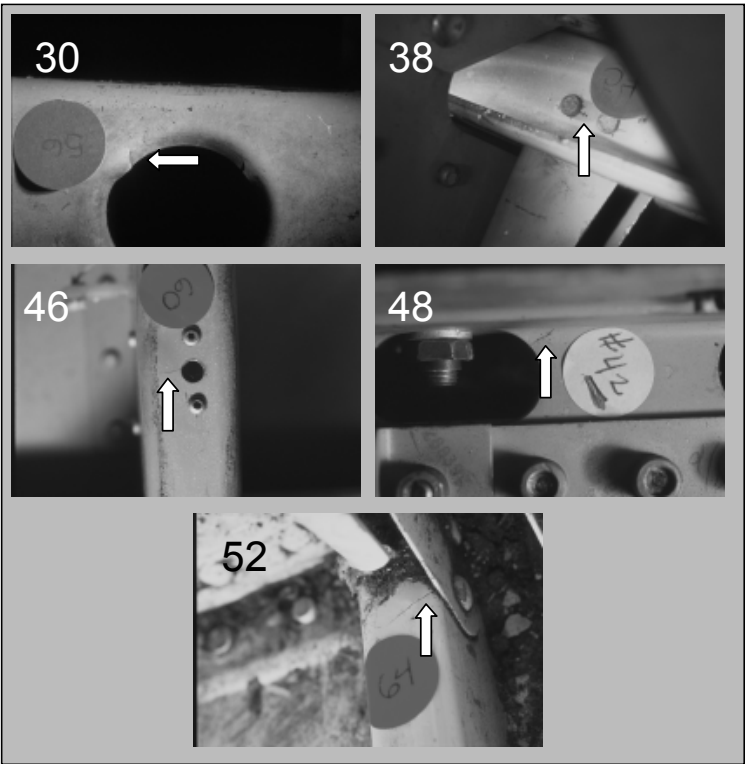
- Ahumada, Jr., A.J. (1996) Simplified vision models for image quality assessment, in SID International Symposium Digest of Technical Papers, ed J. Morreale, Vol. 27, pp. 397-400, SID, Santa Ana, CA.
- Ahumada, A.J., Beard, B.L. (1997). Image discrimination models predict detection in fixed but not random noise. *Journal of the Optical Society of America A* 14, 2471-2476.
- Barten P.G.J. (1999) Contrast sensitivity of the human eye and its effects on image quality. Knegsel: HV Press.
- Ahumada, A.J., Beard, B.L. (1998) A simple vision model for inhomogeneous image quality assessment, in SID Digest of Technical Papers, ed. J. Morreale, vol. 29, Paper 40.1, Santa Ana, CA.
- Beard, B.L., Frank, T.A., Ahumada, A.J. (2003) Using vision modeling to define occupational vision standards, Human Factors and Ergonomics Society Annual Meeting (October 15, Denver, CO).
- Beard, B.L., Hisle, W.A., & Ahumada, A.J. (2002) Occupational vision standards: A review. Report for FAA AAR-100.
- Good, G. W. & Augsburger, A. R. (1987). Uncorrected visual acuity standards for police applicants. *Journal of Police Science and Administration*, 15(1), 18-23.
- Good, G. W., Weaver, J. L., & Augsburger, A. R. (1996). Determination and application of vision standards in industry. *American Journal of Industrial Medicine*, 30(5), 633-630.
- Hellier, C. J. (2001) *Handbook of Nondestructive Evaluation*. New York: McGraw-Hill.
- Padgett, V. R. (1989). Summary of empirical validation of firefighter vision requirements: four field experiments. Paper presented at the International Personnel Management Association Assessment Council 13th Annual Conference held in Orlando, FL. Retrieved July 11, 2002, from <http://home.earthlink.net/~medtox/firevis.html>
- Sheedy, J. E. (1980). Police vision standards. *Journal of Police Science and Administration*, 8(3), 275-285.

APPENDIX A

The next 3 pages show 13 images with 15 cracks used in the study. Crack numbers indicated on each photo correspond to the crack numbers used in the text Figures. Arrows indicate the position of the crack on the image. Numbers and arrows were not present during the experiment.







An Analysis of the Visual Demands Associated with Aviation Maintenance Inspectors

Gregory W. Good¹, Jason J. Nichols¹, Manoj V. Subbaram¹,
Van B. Nakagawara², Ronald W. Montgomery²

¹The Ohio State University College of Optometry, Columbus, OH

²CAMI, FAA, Oklahoma City, OK

Background: Aircraft maintenance inspectors spend many hours searching for defects in aircraft. Vision guidelines exist for nondestructive inspection and testing (NDI/NDT) personnel, but not for visual inspectors. A detailed task analysis is required before job-relevant vision standards can be developed. This study is a descriptive investigation of the visual tasks of visual and NDI/NDT inspectors. **Methods:** Inspectors at aircraft maintenance facilities were observed performing inspections on aircraft and aircraft components. Fixation distances and directions were measured and recorded for inspectors performing visual, fluorescent penetrant, and borescope inspections. Additionally, a visual information survey was completed by 188 inspectors from the different worksites. **Results:** On over 4000 fixations during inspection procedures, near working distances of 50 cm or less were recorded 66.3% of the time. Intermediate distances (>50 cm to 1 m) comprised 23.3% of the fixation distances and were most frequently observed in the performance of borescope and visual inspections. The mean age of inspectors at these locations was 45.1 years. **Conclusions:** The primary duty of visual inspectors is the identification of defects in aircraft when viewed at near and intermediate distances. Data from this study support the need for vision standards for visual inspectors and for as the addition of an intermediate visual acuity requirement to the present distance and near vision standard for all inspectors over 50 years of age.

INTRODUCTION

Maintenance personnel working at aircraft maintenance facilities may have primary responsibilities as visual inspectors where they must use only their vision to assess the condition of aircraft and aircraft components; or they can work in areas where Non-Destructive Inspections (NDI) and Non-Destructive Testing (NDT) are performed. In these workareas, NDI/NDT inspectors often use highly sophisticated imaging and scanning devices (e.g., borescopes, ultrasonic scans, eddy current imaging, X-ray) to aid defect detection. However, even for these inspectors, performing a simple visual inspection is a vital component used to ensure that aircraft are safe to fly. In a recent survey of maintenance facilities, 52% of inspectors were classified solely as visual inspectors, 36% were classified as visual and NDI/NDT inspectors, while only 12% were classified solely as NDI/NDT inspectors (Nakagawara et al., 2003).

Recommended vision standards exist for NDI/NDT personnel (Production and Airworthiness Division, 2001); however, these guidelines do not appear to be based upon a job-task analysis, which documents viewing distances required for efficient task performance. Additionally, no such vision guidelines exist for inspectors who only perform visual inspection tasks. Because of the intimacy between the two inspection classifications (i.e., visual vs. NDI/NDT), most facilities use similar testing requirements for both types of inspectors. While their goals are similar, the two jobs are

inherently different in terms of the visual task and sophistication of testing equipment used.

To the greatest extent possible, vision standards should ensure that workers have the necessary visual skills to perform job-relevant tasks in an efficient and safe manner. For NDI/NDT inspectors, vision skills should be adequate to identify areas of concern (i.e., *detect* potential defects) and to determine if further action is required (i.e., *decide* if a possible defect is within tolerances or if special tests are necessary) (Drury, 2001). Although the NDI/NDT personnel have many tools to aid in the detection of defects (e.g., fluorescent penetrant and magnetic particle inspections, eddy current and ultrasonic devices, borescopes, magnification aids), simple visual inspection may account for up to 80% of all inspections (Goranson and Rogers, 1983).

With advancing age, one gradually loses the normal physiologic ability to focus on near objects. This condition is termed presbyopia. Beginning at age 40, individuals often have difficulty focusing for extended periods at a normal reading distance. For an inspector over 40 years of age, the decline in accommodation may start to affect nearpoint searching. Typically by age 50 almost all focusing ability is lost.

Bifocal lenses can provide appropriate focus for a given working distance, for example, at 16 inches with a +2.5 Diopters (D) reading addition. For a normally-sighted presbyope, with vision correctable to 20/20, these bifocal spectacles would allow for passage of the present Air Transport Association

Specification 105 standard. Should the inspector be required to view at a distance of 32 inches, however, the search area would be 1.25 D out of focus in both the distance and near portions of his spectacles. He / she would now be inspecting the aircraft with reduced visual acuity, estimated to be 20/50 to 20/60. The FAA manages this situation for pilots 50 years of age and over by requiring that pilots see 20/40 or better at both 16 and 32 inches (Nakagawara and Wood, 1998). This age-related requirement is based upon the need for pilots to see cockpit instruments at intermediate distances and the normal physiological changes that limit a person's ability to focus at near and intermediate distances after 50 years of age.

A detailed task analysis with documentation of required working distances is not present in the aviation literature for NDI/NDT and visual inspectors. This study investigated the visual task performed by aviation maintenance inspectors and looks specifically at the viewing distances and directions required to conduct fluorescent penetrant, borescope and visual inspections.

METHODS

The research protocol was approved by the Institutional Review Board of the Ohio State University. Visual and NDI/NDT inspectors at five aircraft maintenance facilities were observed as they performed inspection duties on several types of commercial aircraft (e.g., B727, B737, B767, A320, DC8, DC9, MD80). Various measures of the visual tasks were recorded, along with the specific auxiliary aids used (i.e., flashlight, magnifier, measuring rule), during fluorescent penetrant, borescope, and visual inspection procedures. Additionally, visual inspection tasks were divided into two categories depending upon the major intent of the procedures. These categories were termed "buy-back" and "primary" inspection tasks.

Fluorescent Penetrant Inspections. Fluorescent penetrant inspections (FPI) were observed at only one maintenance facility. Inspections were mainly performed on engine parts. These parts were inspected at the "case" shop or the "rotary" shop, depending on whether the part was a rotating or non-rotating engine component. While good practices for FPI lists 7 moderately independent steps (Drury, 1999), only the inspection (visual detection and decision) portion of the procedure was observed and assessed. Within both shops, engine parts would move along while suspended from an overhead conveyor. Workers would divert individual parts from the main conveyor and move it to their workstations in order to complete the fluorescent penetrant inspection procedure.

Borescope Inspections. Borescope inspections (BI) were observed at 2 of the maintenance centers. The inspection procedure involved using a video borescope to inspect internal engine parts (Drury and Watson, 2000). Inspectors viewed a video monitor as they searched for internal engine defects. At one facility, the engines were separated from the aircraft, while at the other, the engines were inspected while still mounted under the wing.

Visual Buy-Back Inspections. Inspections were termed "buy-back" when inspectors checked jobs individually completed by aviation maintenance technicians (AMTs, i.e., mechanics). These tasks were very specific and generally involved repair or replacement of individual parts or aircraft assemblies. Many involved the inspectors reviewing the AMT's job card for repair descriptions at an inspection station before traveling to the AMT's work bench or aircraft section. A "buy-back" inspection would typically last only 30 to 60 seconds, but could last several minutes when a complicated visual inspection was necessary.

Visual Primary Inspections. Primary inspections were those tasks where workers checked general areas during the initial phases of maintenance to identify specific types of defects identified on work cards. Overall, these inspections could last between several minutes for small jobs to several hours for inspections of large areas.

For FPI, BI, and visual primary inspections, researchers recorded viewing distances and directions at specific points in time while workers performed inspection procedures. Depending upon the type of work and areas under inspection, researchers would record viewing information at 30-second or 1-minute intervals. Therefore, the data represents viewing information similar to that which would be collected if a video recording were sampled at every "nth" frame. For visual buy-back inspections, workers would typically view the indicated parts for only 30 seconds to several minutes. Because of this, only a single fixation distance was recorded for these inspections.

For viewing distance, researchers indicated the distance from the inspector's eyes to the visual target using 7 different distance categories (≤ 33 , 34 to 40, 41 to 50, 51 to 66, 67 to 100, 101 to 200, and > 200 centimeters). These categories represented equal steps in focusing units (i.e., 0.50 Diopters or inverse meters).

For this report, the 7 fixation distance groups were reduced to 3 by merging data from appropriate groups. The fixation distance data in this report are presented as follows:

- a) Near – 50 cm or less,
- b) Intermediate – over 50 cm to 1 meter, and

c) Far – over 1 meter.

For viewing direction data, “up” was marked when the object of regard (OR) was above the level of the inspector’s eyes, “down” was marked when the OR was between eye level and the inspector’s waist, and “full-down” was marked when the OR was below the inspector’s waist.

A Chi Square analysis of the distributions of fixation distance and fixation direction was performed across the three types of inspections (visual, fluorescent penetrant, and borescope).

Finally, a voluntary survey was distributed to visual and NDI/NDT inspectors at the various maintenance facilities that solicited demographic and refractive error correction information (e.g., glasses, contact lenses, refractive surgery).

RESULTS

Data analyzed were from 5 maintenance facilities in the continental United States. Three of these facilities were private, one was a major airline, and one was at a military installation.

Survey. The mean age of inspectors responding to the survey administered at these facilities was 45.1 ± 8.5 years ($n = 188$), and survey responses are summarized in Table 1. Of those responding to the survey (approximately 30% of the entire inspection workforce for these facilities), 49.5% reported wearing spectacles for near work activities, 8.0% reported wearing contact lenses at some time on the job, and 6.9% reported to have undergone refractive surgery. Approximately 30% of the respondents wore no refractive correction at either distance or near. For inspectors over 40 years of age using nearpoint correction, 35% reported wearing single vision lenses, 24% reported wearing traditional bifocals, 35% reported wearing progressive bifocals, 4% reported wearing trifocals, and 2% reported wearing double bifocals. For those wearing contact lenses, 80% reported to wear soft lenses while none of the respondents reported to wear bifocal or monovision contact lenses.

A slight majority of inspectors completing the survey rarely performed any NDI/NDT procedures. Of the respondents, 57.6% reported that less than 10% of their work time is devoted to NDI/NDT procedures. As a group average, however, it was reported that 26.8% of overall inspector time was devoted to NDI/NDT procedures.

Observations. The distribution of fixation distances and directions for visual inspections, fluorescent penetrant inspections, and borescope inspections for over 4,000 recorded fixations are summarized in Table 2.

Fixation Distance. For all inspections, visual detail was often viewed at “normal” reading distances (less than 50 cm for 66.3% of fixations). This was particularly true for fluorescent penetrant inspections where working distances at 50 cm or less were observed over 93% of the time. On the other extreme, however, near fixation distances were observed for borescope inspections 33.4% of the time. For these inspections, borescope inspectors primarily viewed a video monitor positioned at an intermediate distance. Visual inspection tasks were most often performed at near viewing distances (72.2%).

Fixation Direction. With borescope and fluorescent penetrant inspections, fixation direction was mainly confined to normal reading locations (down position). For both inspection types, workers had control of the work environment and could move the visual target to a comfortable position. For visual inspections, workers often had to position their bodies relative to a fixed visual target and, therefore, more variable fixation directions were required. This resulted in viewing up nearly 30% of the time with visual inspections and viewing below the waist nearly 16% of the time. Further analysis showed that for the upward fixations, a vast majority (75%) involved focusing within 50 cm.

Chi-square analysis results across inspection types are shown in Table 2. The distributions for both fixation distance and fixation direction are shown to be different across the 3 inspection methods. Fluorescent penetrant inspection is heavily weighted at the near fixation distance in the normal down position. Borescope inspections are more evenly distributed across all viewing distances but are heavily weighted in the down viewing position. For visual inspections, a wide distribution is found across both fixation distance and direction.

Table 1. Survey Responses.

Inspectors were those respondents that reported to perform NDI/NDT or VI procedures over 50% of their work time.

Outcome	All n = 188	NDI/NDT n = 46	VI n = 103
Mean Age (yrs)	45.1	44.3	45.6
	^a t = 0.67, p = 0.50		
Glasses for Near Inspection	49.5 %	67.4%	42.7%
	^a Chi-Sq = 7.74, p = 0.005		
CL Wearer	8.0%	10.9%	5.8%
	Chi-Sq = 1.18, p = 0.277		
Refractive Surg	6.9%	4.3%	8.7%
	^a Chi-Sq = 0.90, p = 0.344		

^a T-test and Chi-square tests for NDI/NDT and VI comparison.

Table 2. Fixation Distances and Directions (percentages).

Distance	VI	FPI	BS	All
Near	72.2	93.3	33.4	66.3
Intermediate	18.7	6.5	44.7	23.3
Far	9.2	0.2	21.9	10.4
Position	^a Chi-Sq = 620.6, p < 0.001			
Up	29.0	14.2	8.1	17.1
Down	55.4	85.8	88.9	76.7
Full Down	15.7	0.0	2.9	6.2
	^a Chi-Sq = 494.2, p < 0.001			

^aChi-square tests for comparison of 3 type inspections.

DISCUSSION

The establishment of a vision standard shares many similarities with the determination of a cut-off score for any ability test. The essential job functions must be identified as well as the consequences of non-performance. While the frequency of task performance is an important element in setting a standard, task frequency cannot always be equated with task importance. When the consequences of an error are dire (missed crack in a critical component, for example), even a rarely performed task can drive a vision standard. The majority of inspection work performed by all inspectors in this study was performed at viewing distances of less than 50 cm (i.e., 66.3%). Thus, the essence of this work is the identification of defects at near working distances. Coupled with the extreme potential consequences of missing a defect, the frequency data greatly supports the need for a nearpoint visual acuity standard for visual inspectors who are currently not required to meet acuity requirements at any distance.

The data supporting the need for an intermediate visual acuity standard is also strong, especially for visual and borescope inspections. Visual inspectors must observe aircraft components that are difficult to reach and to visualize. These inspectors often cannot physically position themselves to obtain "normal" viewing distances and directions. Intermediate distance viewing is often required. For borescope inspections, workers do have greater control for the inspection. Inspectors can position television monitors for viewing at convenient locations, even though the parts inspected can be relatively inaccessible to the inspector. Borescope inspectors, however, often chose intermediate viewing distances for

viewing the monitor to allow for full body movements to more easily hold and position the borescope probe.

The differences in the distributions of working distances and directions across the different types of inspections are due both to the nature of the inspection task and to the control (or lack of control) the inspector has on the part being inspected. With FPI, the majority of the work is done at near working distances in a normal reading position (down). This was the case for fluorescent penetrant because most inspections are done on individual parts taken off aircraft, allowing greater control of part positioning.

Visual inspectors have the least viewing flexibility as the object of regard is often firmly fixed to the aircraft and inspectors must change body and head position, often in cramped quarters, to gain an acceptable viewing posture. Nearly 20% of visual inspections are done at an intermediate viewing distance (between 50 cm and 1 meter). Visual inspectors often inspect large areas of an aircraft for cracks and other defects from intermediate distances. Because a longer working distance translates into smaller visual angles for visual detail subtended to the eye, it could be argued that it is more important for inspectors to be capable of clear focusing at intermediate distances than it is for near working distances. For borescope inspections, nearly one-half (44.7%) of the viewing distances were observed to be between 50 cm and 1 meter. It is clear that a large portion of aircraft inspection must be done with a fixation distance of greater than 50 cm.

Because of our normal physiologic accommodative ability, if a worker under 40 years of age can pass a vision standard at a given distance using normal, single vision glasses, he/she should be able to pass the same standard at all working distances. For workers older than 50 years, however, specially designed multifocal lenses may be required to allow sharp vision at intermediate and near working distances.

As the mean age of surveyed inspectors is 45.1 years, a large proportion of inspectors have lost significant natural accommodative power. Eyewear must be designed with viewing distances and directions in mind. Although the majority of fixation directions for aircraft inspection correspond to the normal bifocal position (slightly down), much visual inspection activity is directed upward (29.0%) and at intermediate to long viewing distances (27.9%). Inspectors should thoroughly discuss the variations in object distance

and direction required of their jobs with their eye care practitioners. In order to ensure clear and comfortable vision at all working distances, special eyewear designs may be required. Inspectors older than 50 years may require trifocals or progressive addition bifocals (i.e., no-line) to allow clear vision at all required viewing distances. As working distances vary regardless of the viewing directions, it may be beneficial to use clip-on near lenses to accommodate some working distances and/or awkward directions. A set of clip-on lenses of different powers can be obtained to ensure that clear focusing is obtainable at all fixation distances.

The data presented supports vision requirements for visual inspectors as well as the addition of an intermediate visual acuity requirement to the present distance and near vision standard for all inspectors over 50 years of age. As inspectors age, more frequent vision screenings would help ensure that refractive correction is adequate to accommodate the three working distances. It is impossible to design eyewear, however, that will allow all fixation directions and head positions to be capable of clear vision at distance, intermediate and near distances. Therefore, a worker education program should be included within the overall vision program. Such a program will help inspectors understand the limitations of multifocal lenses for aviation inspection tasks and learn what lens devices are available to better accomplish their visual tasks in a safe and efficient manner.

REFERENCES

- Drury CG. (1999). Human Factors Good Practices in Fluorescent Penetrant Inspection. Human Factors in Aviation Maintenance - Phase Nine, Progress Report, FAA/Human Factors in Aviation Maintenance, via <http://hfskyway.faa.gov>.
- Drury CG. (2001). Good Practices in Visual Inspection. Human Factors in Aviation Maintenance - Phase Nine, Progress Report, FAA/Human Factors in Aviation Maintenance, via <http://hfskyway.faa.gov>.
- Drury CG, Watson J. (2000). Human Factors Good Practices in Borescope Inspection. FAA/Office of Aviation Medicine, Washington, D.C., via <http://hfskyway.faa.gov>.
- Goranson UF, Rogers JT. (1983). Elements of damage tolerance verification, 12th *Symposium of International Commercial Aeronautical Fatigue*, Toulouse, France.
- Nakagawara VB, Montgomery RW, Good GW, Nichols JJ (2003). A demographic profile of nondestructive inspection and testing (NDI/NDT) personnel: A preliminary report. Federal Aviation Administration, Office of the Chief Scientist for Human Factors (AAR-100), FY03 Aviation Maintenance, General Aviation, and Vertical Flight. University of Nevada Reno, 10-12 September 2003.
- Nakagawara VB, Wood KJ. (1998). Clinical application of new civil airman vision standards and certification procedures. *Journal of the American Optometric Association*; 69:144-50.
- Production and Airworthiness Division (AIR-200), FAA (2001). *Qualification Standards for Nondestructive Testing (NDT) Inspection/Evaluation Personnel*, via http://www.faa.gov/certification/aircraft/NDTQ_UALSTDMEMO.htm.

Visual Functioning of Aviation Maintenance Inspectors

Gregory W. Good¹, Jason J. Nichols¹, Manoj V. Subbaram¹, Van B. Nakagawara²,
Ronald W. Montgomery², Ross Collins¹

¹The Ohio State University College of Optometry, Columbus, OH

²CAMI, FAA, Oklahoma City, OK

Background: While an FAA vision standard does not exist for aviation maintenance inspectors, guidance has been given for vision recommendations for NDI/NDT inspectors. This study measured the vision functioning of 150 volunteer inspectors to determine if inspectors' vision complies with the recommended vision standards and met the occupational demands of aircraft maintenance inspections. **Methods:** Vision tests were performed on NDI/NDT and visual inspectors at 2 aircraft maintenance facilities. The results were compiled and compared to the recommended vision standards for NDI/NDT personnel. **Results:** Mean age of inspectors was found to be 44.7 years. Mean visual acuity of inspectors was 20/16.6 and 20/16.8 at 16 foot and 16 inch test distances, respectively. All inspectors met the present distance VA recommendations and only one failed to meet the present recommendation for near vision. **Conclusions:** Aircraft maintenance inspectors have excellent vision functioning. This indicates that medical personnel at these maintenance facilities are adequately screening employees. In spite of these excellent results, inspectors should be educated on the limitations of focusing that accompanies aging and provided with various focusing devices to allow clear and comfortable vision at all required viewing distances and directions.

INTRODUCTION

Inspection tasks for aircraft maintenance inspectors are visually demanding. Whether personnel conduct inspections using only a flashlight, a simple magnifier, or sophisticated NDI/NDT equipment, visual identification is the primary method used to find cracks and other defects which affect the integrity of an aircraft or aircraft component. The National Transportation Safety Board (NTSB) has identified the failure by inspectors to detect visible corrosion, cracks, or inclusions as a causative factor of several aviation accidents (NTSB, 1998, 1990, 1989).

In spite of the importance of vision to the inspection process, the Federal Aviation Administration (FAA) had not required maintenance inspectors to meet occupational vision standards. The FAA's Production and Airworthiness Division (AIR-200) recognized this lack of national policy and prepared a memorandum (Production and Airworthiness Division, 2001), dated September 26, 2001, to address the issue. This memorandum follows an FAA advisory circular from February 1999 that addressed the same topic (FAA, 1999). Several national and international organizations have put forth recommendations for qualifications of NDI/NDT personnel that include initial and recurrent training, levels of competence, and minimum vision standards and test intervals. The September 2001 FAA memorandum identified the

standards thought to be acceptable for assuring that only qualified individuals perform NDI/NDT inspections and procedures:

- 1) MIL-STD-410E, Military Standard
- 2) ATA Specification 105, Air Transport Association
- 3) AIA-NAS-410, Aerospace Industries Association
- 4) ISO 9712, International Standards Organization

The memorandum further describes the generic elements of the different standards and states minimal requirements that organizations developing NDI/NDT qualification programs should meet. In terms of vision testing, the memorandum is summarized below:

1. Vision Examinations: NDT personnel should receive documented vision and color blindness testing at reasonable intervals (one to two years, shorter preferred). Vision examinations shall be administered by personnel in accordance with the standard to determine qualification.

(a) Near Distance Vision Requirements:

Natural or corrected near distance acuity in at least one eye capable of reading the Jaeger #1 Test Chart or equivalent at a distance of not less than 30 cm.

(b) Color Vision Requirements:

Ability to differentiate among colors used in NDT method(s).

These vision guidelines are specifically written for NDI/NDT personnel and lack the specificity required to ensure uniformity of compliance throughout the industry. No such guidelines exist for visual inspectors. Because of the intimacy between the two inspection classifications (i.e., visual inspection vs. NDI/NDT), however, most maintenance facilities use similar testing requirements for both types of inspectors.

In terms of visual acuity, the ATA Specification 105 standard includes a distant visual acuity measure, albeit lenient (20/50), while the AIA-NAS-410 and FAA guidance memorandum do not. In addition, other vision requirements set forth in various industry programs are not uniform. The training manual for NDI/NDT personnel for one national airline lists visual acuity requirements at nearpoint of 20/25 in at least one eye and at distance of 20/30. At another airline, the requirements are more rigorous with a nearpoint requirement of 20/20 and a distance requirement of 20/25. Additionally, the question of an intermediate distance visual acuity requirement is not addressed within any of the aforementioned documents, even though inspectors performing NDI/NDT procedures frequently use working distances between 16 and 80 inches.

In this study, the on-the-job visual capabilities of 150 representative visual and NDI/NDT inspectors were measured. The intent was to determine the visual status of a representative group of inspectors to predict what effect, if any, a change in the present vision standards would have on the present workforce. It is also hoped that visual and medical information obtained can help determine if the present recommendation for the frequency of vision assessment (i.e., not greater than every 2 years) is adequate to ensure a visually competent workforce. The results could also be used to determine whether the present medical surveillance programs employed at the subject facilities are adequately ensuring that inspectors meet the current vision guidelines.

METHODS

The research protocol was approved by the Institutional Review Board of the Ohio State University. Vision screening was performed at two aircraft maintenance facilities. Facility #1 was a private maintenance facility, while facility #2 was

a national airline. Various vision measures were taken on 150 volunteer visual and NDI/NDT inspectors (59 at facility #1 and 91 at facility #2). After a short visual and medical history that included documentation of age, experience as an inspector, and whether vision care insurance was present, subjects underwent the following visual tests with their current corrections (if appropriate):

- a) **Distance Visual Acuity in each eye (LogMAR chart),**
- b) Distance Binocular Low Contrast Visual Acuity (Bailey-Lovie Chart),
- c) Binocular Visual Acuity at 32 inches,
- d) Binocular Visual Acuity at 16 inches,
- e) Global and Local Nearpoint Stereoacuity,
- f) Color Vision (Ishihara Pseudoisochromatic Plates (PIP) and Farnsworth D-15 for PIP failures),
- g) Nearpoint Contrast Sensitivity (Pelli- Robson Chart), and
- h) Intraocular Pressure (Tonopen).

An objective measure of refractive error was also taken (i.e., autorefractor); however, the results of that testing are not reported here. Additionally, the powers of the current spectacles were measured and lens designs were recorded (i.e., normal bifocal, multifocal, occupational bifocal, or single vision lenses). Measures of vision were taken by experienced eyecare personnel from The Ohio State University College of Optometry and the Vision Research Team of the Civil Aerospace Medical Institute (CAMI) from the FAA in Oklahoma City, Oklahoma.

RESULTS

The results of the screening are presented in the appropriate sections below and are divided into classes for the different types of inspections (i.e., visual vs. NDI/NDT). A summary of these results are presented in Table 1.

Inspector Demographics.

Inspector Age. The mean age of these 150 inspectors is 44.6 years \pm 7.9 years. The ages did not differ significantly between examination sites (Two Sample T, t-value = -0.93, p = 0.357). The age of visual and NDI/NDT inspectors were documented in an associated study (Good et al., 2004) for 183 inspectors from 5 maintenance facilities. One of those facilities was Facility #1 from the present study. Therefore, only the ages of inspectors from facility #2 were compared to the

previously surveyed population. The ages of these populations also did not differ significantly (Two Sample T, t-value = -0.89, p = 0.375). Therefore, the two populations were combined to give an age value representative of the overall inspector population. The figures for the combined population (n = 274) are:

Mean Age = 44.8 ± 8.4 years,
Range 25 to 68 years.

Systemic and Ocular Disease. Only 6 inspectors, or 4% of our subject population, reported having diabetes. This is less than the reported 8.7% of the US population over 20 years of age with diabetes (Cowie et al., 2004). Twenty-seven participants (18% of our subjects) reported having high blood pressure. Estimates in the United States are that 31.3% of adults have high blood pressure (Fields et al., 2004). Using these systemic conditions as overall health indicators, one could infer that the subject population was more healthy than the US population as a whole.

Only 2 subjects (1.3%) reported being treated for glaucoma. National estimates are that 1.9% of the over 40 population has glaucoma (Prevent Blindness America / National Eye Institute, 2002). As many of the participants were under 40 (26.7%) and the prevalence of glaucoma increases with age, it was not possible to compare our figures to the over 40 national estimate.

Experience and Classification of Inspection. Study participants were classified as either visual or NDI/NDT inspectors based upon which activity occupied the majority of their work time. Fifty of the participants reported that NDI/NDT inspections accounted for more than 50% of their workdays (33.8% classified as NDI/NDT inspectors), while 98 reported less than 50% (66.2% classified as visual inspectors). Two participants reported an equal, 50/50 split of work activities. Data from facility #2 were then combined with the previously reported survey data (Good et al., 2004). The number of years of aviation inspection experience for surveyed inspectors did not differ significantly between inspector classification (visual versus NDI/NDT, t-value = 0.21, p = 0.836) in spite of visual inspectors being slightly older (mean age 45.7 years [visual] to 43.3 years [NDI/NDT], t-value = 2.02, p = 0.045). The inspector experience for the combined populations (n = 274) is:

Mean Years as Inspector = 10.3 ± 7.7 years,
Range: < 1 year to 42 years.

Table 1
Screening Results
Mean \pm SD

Inspector Age	
Facility #1	45.3 ± 7.2 years
Facility #2	44.1 ± 8.3 years
Overall	44.6 ± 7.9 years
Visual Acuity (with correction)	
(Log MAR, 20/20 = 0.0)	
16 ft. (better eye)	-0.08 ± 0.08 (20/16.6)
32 in (binocular)	-0.17 ± 0.09 (20/13.4)
16 in (binocular)	-0.08 ± 0.05 (20/16.8)
Contrast Sensitivity	
Low Contrast VA (16 ft)	
LogMAR	0.03 ± 0.09 (20/23.2)
Pelli-Robson (1 m)	1.93 ± 0.05
Stereopsis (seconds of arc)	
Global	255.0 ± 45.5
Local	33.2 ± 35.1
Intraocular Pressure	
Tonopen	13.7 ± 3.3 mm Hg
Color Vision (% Failed)	
Ishihara PIP	3.3%
Farnsworth D-15	2.0%

Visual Measures.

Visual Acuity. Visual acuity measures were taken with correction (if normally worn by the inspector) for each eye at 16 feet (distance), while acuity for near (16 inches) and intermediate (32 inches) distances was measured binocularly. At distance, the mean visual acuity of the better eye was better than 20/16.6; and, only 9 of the 150 inspectors had less than 20/20 with none measuring worse than the 20/50 specified by the ATA specification 105 recommendation. At nearpoint the mean visual acuity was 20/16.8. Eleven individuals scored less than 20/20, but only 1 failed (by a single letter) to meet the 20/25 ATA recommendation. Although ATA specification 105 does not specify an intermediate visual acuity requirement, visual acuity at 32 inches was found to be outstanding (mean acuity = 20/13.4). Only 5 individuals failed to see 20/20 at the intermediate distance.

Contrast Sensitivity. Pelli-Robson contrast sensitivity measures were excellent for these inspectors. Only a single inspector had contrast sensitivity below 1.80. The mean contrast sensitivity was 1.93 (contrast threshold = 1.17%). Low contrast visual acuity (LCVA) is a test, which incorporates elements of both contrast sensitivity and visual acuity. It is often claimed to be a better

indicator than high contrast visual acuity of “real-world” performance. Of the 150 inspectors, 145 had LCVA measured at distance of 20/32 or better. The mean LCVA was 20/23.2.

Stereoacuity. Nearpoint stereoacuity was measured using the Randot Stereo Test. Measures of both local and global stereopsis were made. For local stereopsis a median value of 20” of arc was found (mean = 33.2” of arc). This is the limiting value for the test. Only 2 of the 150 inspectors had worse than 70” of arc on this test. For global stereopsis only 1 inspector was unable to identify any target and only 4 additional inspectors measured less than the best possible.

Color Vision. Five of the 150 inspectors (3.3%) were found to have abnormal color vision by failing the Ishihara PIP test. Of these five, three showed a moderate to severe color vision defect by failing the Farnsworth D-15 test.

Intraocular Pressure. Intraocular pressure (IOP) was measured using the Tonopen tonometer. Mean intraocular pressure as 13.7 mm Hg. Only one inspector was found with IOP measures above 21 mm Hg.

Refractive Correction. For the 150 inspectors, eighty-nine wore some type of spectacles. Sixty-six required a special correction for near activities (see Figure 1). Of these, 25 were single-vision near glasses, 32 were no-line, progressive bifocals, and only 9 were traditional straight-top, line bifocals. None of these workers wore special design, occupational multifocals with a near focusing segment across the top of the lenses.

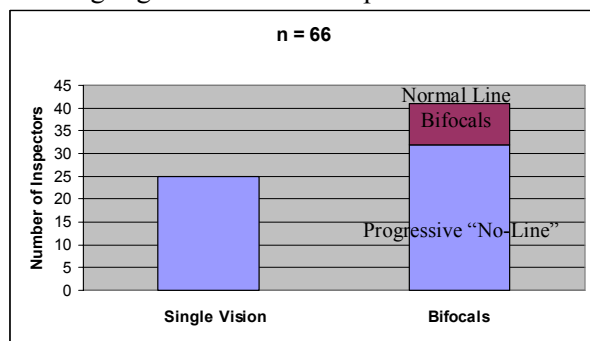


Figure 1. Design of Nearpoint Spectacles

DISCUSSION

In terms of overall systemic and ocular health, the inspectors participating in this study appear to be healthier as a group than the overall US population. As the job duties of most inspectors require a good deal of physical exertion, it is not unreasonable to assume that their active workdays are positive factors in this finding.

The visual functioning of the 150 inspectors examined in this study was also excellent. Only 9 inspectors had less than 20/20 visual acuity at distance with the better eye, and none failed to meet the 20/50 ATA specification 105 distance visual acuity recommendation. At nearpoint, only 11 inspectors had less than 20/20 visual acuity and only 1 did not meet the 20/25 requirement, and this was by just a single letter.

These inspectors also demonstrated excellent visual acuity at the intermediate distance. This was expected for those inspectors at 45 years of age or less. Although near focusing ability decreases with age, those 45 and younger should be able to focus for short periods to near 16 inches from the eyes. For the inspectors older than 45, reading glasses or bifocals become a requirement to focus near objects; and, stronger bifocals can focus for very near objects but leave individuals a focusing “dead-zone” at intermediate distances, where details are blurry through both the distance and bifocal portions of glasses. Individuals older than 50 years must allow for this in the design of their glasses when objects at intermediate distances are viewed. The FAA recognizes this eventuality by requiring pilots with Class I and II medical certificates to demonstrate relatively sharp visual acuity at 32 inches (Nakagawara and Wood, 1998).

The older inspectors in this study were able to see clearly at intermediate distances largely from having progressive addition (i.e., no-line) bifocals. The powers in progressive lenses gradually change from the top to the bottom portions for the lenses allowing clear focusing at intermediate distances. Seventy-eight percent of the bifocal wearers in this study used progressive addition lenses. Of concern, however, is focusing on objects positioned off to the side or superiorly in the field of view. This is often the case for visual inspectors. It becomes difficult to position the head to see through the inferiorly placed bifocal segment when the object of interest is off to one side or above the head. Special care must be taken in the design of eyewear to ensure clear, comfortable vision for these positions.

Inspectors should discuss spectacle design options with their eyecare providers. Clip-on near focusing lenses, occupational bifocals, and special designed lenses to be used solely for inspection tasks are 3 lens alternatives that can provide in focus imagery for all distances and directions required during inspection tasks.

In conclusion, inspectors at the two facilities where testing was conducted appear to have adequate vision function to effectively perform their responsibilities. However, since the subject selection process was entirely voluntary, results could vary for the inspector population as a whole. Proper vision testing at appropriate intervals is the key to maintaining a visually healthy workforce. The addition of an age-related intermediate visual acuity requirement and guidance for selection of appropriate refractive correction would provide additional safeguards to ensure that inspectors retain optimal vision performance as they age.

REFERENCES

- Cowie CC, Rust KF, Byrd-Holt D, et al. (2003). Prevalence of diabetes and impaired fasting glucose in adults – United States, 1999-2000. *MMWR* 52(35):833-837.
- Federal Aviation Administration (1999). *Training Qualification Certification of Nondestructive Inspection (NDI)*. Advisory Circular ACNO. 65-XX via <http://hfskyway.faa.gov>.
- Fields LE, Burt VL, Cutler JA, Hughes J, Roccella EJ, Sorlie P. (2004).. The Burden of Adult Hypertension in the United States 1999 to 2000, A Rising Tide. *Hypertension*; 44: 1-7.
- Good G, Nichols J, Subbaram M, Nakagawara V, Montgomery R. (2004). An analysis of Visual Demands of Aviation Maintenance Inspectors. Federal Aviation Administration, Office of the Chief Scientist for Human Factors (AAR-100), FY04 Aviation Maintenance, General Aviation, and Vertical Flight. Washington, DC November 16, 2004.
- Nakagawara VB, Wood KJ. (1998). Clinical application of new civil airman vision standards and certification procedures. *Journal of the American Optometric Association*; 69:144-50.
- National Transportation Safety Board (1998). Uncontained engine failure. Delta Airlines flight 1288. McDonnell Douglas MD-88, N927DA, Pensacola Regional Airport, Pensacola, July 6, 1996. NTSB/AAR-98/01.
- National Transportation Safety Board (1990). Aircraft Accident Report. United Airlines Flight 232, McDonnell Douglas DC-10-10, Sioux Gateway Airport, Sioux City, July 19, 1989. NTSB/AAR-90/06.
- National Transportation Safety Board (1989). Aircraft Accident Report. Aloha Airlines, Flight 243. Boeing 737-200, N73711, near Maui, Hawaii, April 28, 1988. NTSB/AAR-89/03.
- Prevent Blindness America / National Eye Institute. (2002). Vision Problems in the US, via <http://www.preventblindness.org>.
- Production and Airworthiness Division (AIR-200), FAA (2001). *Qualification Standards for Nondestructive Testing (NDT) Inspection/ Evaluation Personnel*, via <http://www.faa.gov/certification/aircraft/NDTQUALSTDMEMO.htm>.

YEAR1: DEVELOPMENT OF A WEB-BASED SURVEILLANCE AND AUDITING TOOL TO REDUCE AVIATION MAINTENANCE ERRORS (WebSAT)

Kunal Kapoor, Pallavi Dharwada, Nikhil Iyengar, Joel S. Greenstein,
Anand K. Gramopadhye
Human Computer Systems Laboratory, Clemson University
Clemson, South Carolina

The safety and reliability of air transportation depends on minimizing inspection and maintenance errors that occur in the aircraft maintenance system. Efforts have been invested to track maintenance errors. These efforts are reactive in nature: they analyze maintenance errors after their occurrence. There is a lack of standardization in the assessment of maintenance errors across the maintenance industry. Surveillance and auditing of maintenance activities are two important functions which help ensure airworthiness of an aircraft. A system that will document the processes and outcomes of these maintenance activities and will make this documentation more accessible will accomplish the goal of this research to reduce maintenance error. Such a system would then support robust and safer aircraft maintenance operations. Our research is developing a web-based surveillance and auditing tool (WebSAT) that promotes a standardized format for maintenance data collection, reduction and analysis to proactively identify the factors contributing to improper maintenance.

INTRODUCTION

The aircraft maintenance system is complicated (Gramopadhye, Drury and Prabhu, 1997), with interrelated human and machine components. Realizing this, the FAA has pursued human factors research for some time now under the National Plan for Aviation Human Factors (FAA, 1991; FAA, 1993) to fulfill the mission of the FAA's Flight Standards Service of promoting safety by setting certification standards for air carriers, commercial operators, air agencies, and airmen.

A study conducted by Boeing and the US Air Transport Association (1995) found that maintenance error was a crucial factor in aircraft accidents from 1982 to 1991, contributing to 15% of the commercial hull loss accidents where five or more people were killed. Rankin and Allen (1995) established the economic costs of these maintenance errors, estimating that 20 to 30% of in-flight shutdowns are due to maintenance error, 50% of flight delays are due to engine problems caused by maintenance errors, and 50% of flight cancellations are due to engine problems caused by maintenance errors. The need is apparent for a proactive system which will help track maintenance errors, identifying both potential problem areas and the factors causing errors. If such a system is developed it will be possible to better manage maintenance errors, resulting in aircraft maintenance which is safer and more robust.

Problem Statement

To minimize maintenance errors, the aviation maintenance industry has developed methodologies to investigate maintenance errors. The literature of human error is rich, having its foundations in early studies analyzing human error made by pilots (Fitts and Jones, 1947), human error work following the Three Mile Island accident, and recent research in human reliability and the development of error taxonomies (Norman, 1981; Rasmussen, 1982; Reason, 1990; Rouse and Rouse, 1983; Swain and Guttman, 1983). This research has centered on analyzing maintenance

accidents and incidents, a recent example being the Maintenance Error Decision Aid (MEDA) (Rankin, Hibit, Allen and Sargent, 2000). This tool, developed by Boeing along with representatives from British Airways, Continental Airlines, United Airlines, the International Association of Machinists and the US Federal Aviation Administration, helps analysts identify the contributing factors leading to an accident. Various airlines have developed internal procedures to track maintenance errors. One such methodology is the failure modes and effects analysis approach (Hobbs and Williamson, 2001) that classifies potential errors by expanding each step of a task analysis into sub-steps and then listing the potential failure modes. The US Naval Safety Center developed the Human Factors Analysis and Classification System- Maintenance Extension Taxonomy and the follow-up web-based maintenance error information management system to analyze naval aviation mishaps (Schmidt, Schmorow and Hardee, 1998; Shappell and Wiegman, 1997, 2001) and later used to analyze commercial aviation accidents (Wiegman and Shappell, 2001). Although valuable in terms of their insights into performance-shaping factors leading to maintenance errors following their occurrence, these efforts are reactive in nature. Maintenance error tracking efforts are also ad hoc in nature, varying across the industry with little standardization. The lack of standardization in data collection, reduction and analysis is the single biggest drawback in the analysis of maintenance errors within and across the maintenance industry. This research is developing a web-based surveillance and auditing tool (WebSAT) that promotes standardized data collection and analysis. Surveillance, auditing, and airworthiness directives are the activities which will be the primary data sources for WebSAT, as shown in Figure 1.

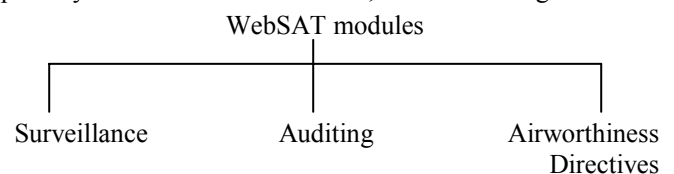


Figure 1. Data sources for WebSAT

Substantial maintenance vendor and fuel vendor surveillance activities will form the basis for our inputs on surveillance activities. Technical audits, internal audits, self audits, and fuel, maintenance and ramp audits will form the basis for inputs on auditing activities. Airworthiness directives data will be derived from work instruction cards and engineering orders. For the purpose of illustration, we use surveillance activity as an example to describe our initial development efforts in this paper.

Surveillance: Surveillance is the day-to-day oversight and evaluation of the work contracted to an airframe substantial maintenance vendor or fuel vendor to determine the level of compliance with the airline's Continuous Airworthiness Maintenance Program (CAMP) and General

Maintenance Manual (GMM). The objective of surveillance is to provide the airline, through the accomplishment of a variety of specific surveillance activities on a planned and random sampling basis, an accurate, real-time, and comprehensive evaluation of how well each maintenance vendor is complying with airline and FAA approved policies and regulatory requirements. WebSAT will perform surveillance activities to ensure that a consistent level of supervision is maintained over maintenance and inspection operations. The system will seek input from various sources, including In-Process Surveillance, Verification Surveillance, Final Walk Around, Aircraft Walk Around, Inspection, Storage, among others, as shown in Figure 2.

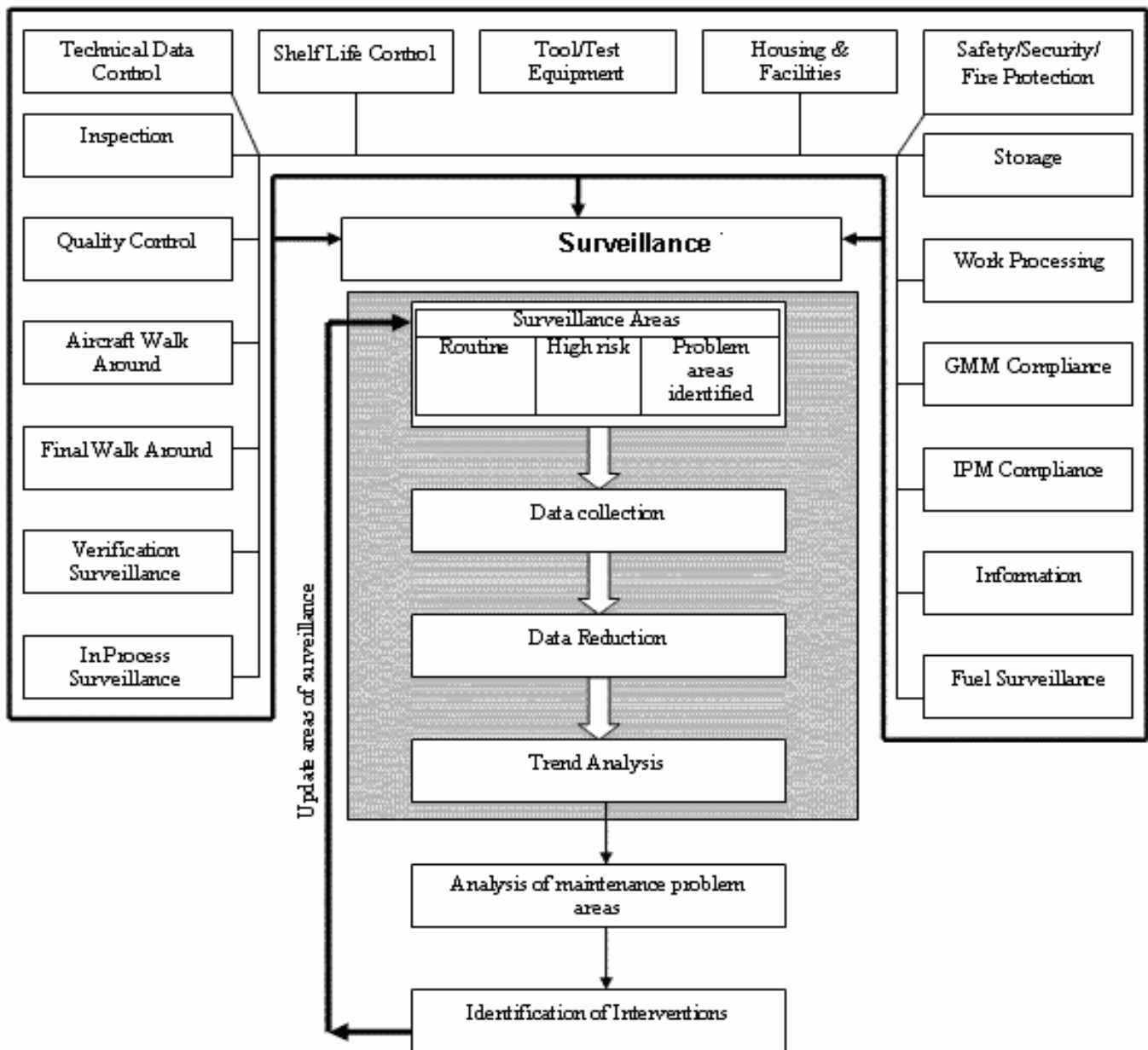


Figure 2. Data sources involved in a surveillance activity

These are the sources which provide the most information about maintenance and inspection errors and hence are termed the potential process measures that affect the performance of the surveillance activity. Similar variables are being identified for the other activities mentioned in Figure 1, namely auditing and airworthiness directives.

Data collected from these diverse sources will be analyzed to identify potential problem areas. The identification of these problem areas will let the industry prioritize factors that transcend the individual airlines to systematically reduce or eliminate potential errors. The WebSAT system is being developed with a specific aviation partner (FedEx in Memphis, TN) to ensure the needs of the aviation community are addressed. It will be made available as an application that can be downloaded for use by each maintenance facility.

METHODOLOGY

The research is being conducted in three phases.

Phase 1: Identification of Process measures and Data Sources

- *Data Gathering Techniques:* Based on the various factors that influence the choice of a data gathering method (Iyengar et al., 2004), interviews, observation sessions, document study and questionnaires were adopted as techniques to gather data on the processes of surveillance, auditing, and airworthiness directives so as to accomplish the task of identification of process measures.

- *Process Measures:* To achieve standardization in data collection, data needs to be collected on certain variables which measure maintenance processes and eliminate existing inconsistencies. These variables are defined by the research team as process measures.

The process measures incorporate the response and observation-based data collected during surveillance, audits, and the airworthiness directives control processes. Once data is captured in terms of these process measures, data analysis can be conducted to identify the potential problematic areas affecting the safety of an aircraft. In this stage of data analysis, the performance of processes and those conducting these processes will also be evaluated.

Process measures for surveillance, auditing and airworthiness directives work functions were identified by the research team based on human-factor principles, utility of data being captured, and working around mental models of quality assurance personnel.

- *Validation:* In order to ensure that the identified process measures are representative of those used by most maintenance entities, an online survey has been conducted with the partnering airlines. This survey has been conducted in two stages. Since, the data collected to identify the process measures was from the industry partner FedEx, the research team conducted an initial survey with the users of FedEx in stage 1 and used the findings from this survey to refine the process measures and thereby the final survey before sending

it to the other airlines. The results from the second stage of this survey are still awaited.

- *Finalize the list of process measures:* Based on the results of the survey, the research team will identify the limitations in using the specific process measures identified and finalize the list of process measures that can be used in collecting data from various maintenance processes.

The first phase of the research will finalize the list of process measures.

Accomplishments from Phase I

- Identified the *process measures* using FedEx's existing methodologies, desktop procedure manuals, C.A.S.E standards and human factors guidelines.
- Developed an online survey to validate the identified process measures with research partners.
- Conducted a survey with FedEx personnel to validate the identified process measures.
- Refined the survey based on the input from FedEx and sent to other partnering airlines.

Phase 2: Develop Prototype of Surveillance and Auditing Tool

- *Product phase:* The research team has come up with the project mission statement as shown in Table 1, specifying the vision for the product, the target market, the project goals, the key assumptions, the constraints, and the stakeholders.

- *Needs analysis phase:* The data that was gathered from interviews and observation sessions was used by the research team to identify customer needs, and establish the relative importance of the needs.

- *Product specifications phase:* In the subsequent stage, the researchers will develop a preliminary set of target specifications.

- *Concept generation and selection phase:* After developing the target specifications for the product performance, the team will generate concepts for developing the product and will select the most promising one to carry out the tool development.

- *Detail design of selected concept to create an initial working prototype:* During this phase, low fidelity prototypes which incorporate detailed design are developed and further working prototypes are developed iteratively to conduct testing with the users.

- *Testing and refinement of the initial working prototype is carried out with representative users in the next stage.*

The second phase of the research will deliver a refined prototype to FedEx for trial use.

Table 1. Mission Statement of WebSAT

Mission Statement: Web-based Surveillance and Auditing Tool Prototype	
Product Description	<ul style="list-style-type: none"> An application, incorporating a recommended categorization and data collection scheme for maintenance auditing and surveillance application; a data reduction module that allows the analysts to conduct central tendency analysis and data analysis module that facilitates trend analysis.
Key Business Goals	<ul style="list-style-type: none"> Achieve standardized data collection/reduction and analysis of maintenance errors across the geographically dispersed entities of the airline industry Develop a proactive system that captures maintenance errors Generate trend analysis
Primary Market	<ul style="list-style-type: none"> FedEx
Secondary Market	<ul style="list-style-type: none"> Other airlines in the Airline Industry
Assumptions & Constraints	<ul style="list-style-type: none"> SQL server, ASP.NET
Stakeholders	<ul style="list-style-type: none"> FedEx QA Department Airworthiness Directives Control Group Other airlines

WebSAT Research Framework:

The WebSAT research framework shown in Figure 3 has 3 tiers associated with it. In tier 1, relevant data collected from the three modules (surveillance, auditing and airworthiness directives) will be analyzed using the identified process measures which allow us to evaluate the effectiveness of each module.

Further analysis of data will lead us to the categories in tier-2 which evaluates the performance of the airline across the three modules. These categories are factors such as cost, economy, etc. which have a direct bearing on the impact on the safety of an airline.

Our research team will then conduct analysis of tier-2 and estimate safety index of the airline by identifying the risk-causing factors represented in tier-3. In tier 3 it is

demonstrated that the variables are of 2 kinds: risk and non-risk. The upper management is interested in the risk or impact variables, which will be indicated by the tool. The research team finds it appropriate to report results of analysis for non-risk variables, contemplating that useful input will be generated.

Phase 3: Develop Data Analysis and Validation Module

- Develop advanced data analysis tools that include multivariate analysis and risk assessment.
- Validate using field data.

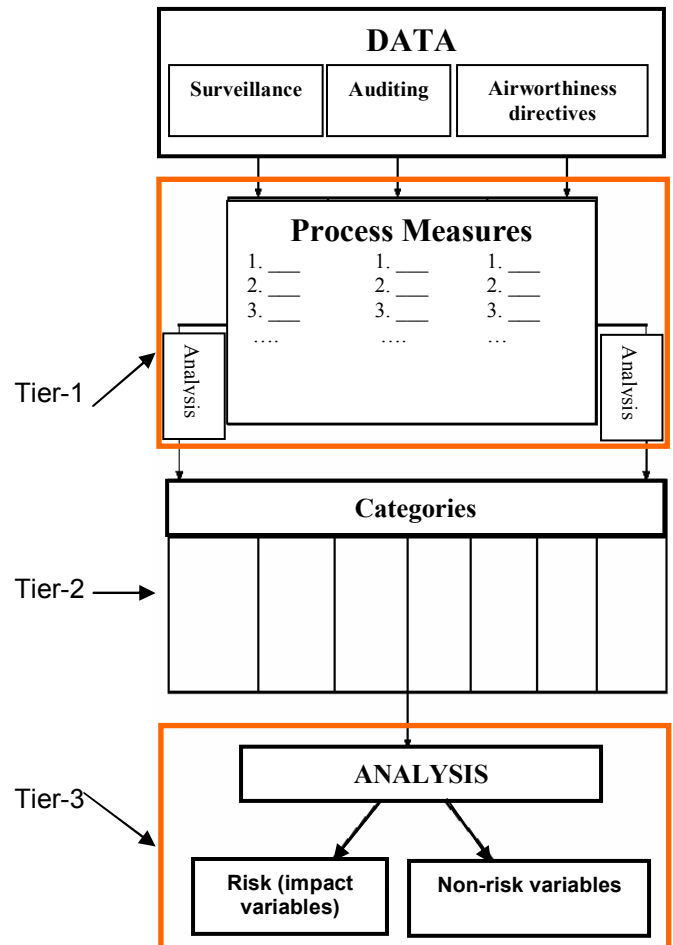


Figure 3. WebSAT Framework Prototype

SIGNIFICANCE AND IMPACT OF WEBSAT

The development of a web-based surveillance and auditing tool has the potential to reduce maintenance errors impacting aviation safety. The specific advantages of this tool are the following: (1) a proactive approach reduces maintenance errors by identifying problem areas and error contributing factors; (2) the adoption of this tool by the aircraft maintenance industry promotes standardization in collection, reduction and analysis of maintenance error data; (3) this standardization will result in superior trend analysis of problem areas; and (4) the findings can be shared by manufacturers, airlines, repair stations and air cargo handlers

to identify and prioritize factors which lead to maintenance errors.

CONCLUSION

In summary, the objective of this research is to: (1) identify an exhaustive list of process measures that affect aviation safety and transcend various aircraft maintenance organizations; (2) design and develop web-based surveillance and auditing tool which uses the identified set of process measures for data analysis. The results of this research will be disseminated to the aviation community via a number of avenues. These include scholastic publications and training software available for download from the FAA's web site and the regular communication of the results of this research to industry partners.

ACKNOWLEDGEMENTS

This research is supported by a contract to Dr. Anand K. Gramopadhye and Dr. Joel S. Greenstein, Department of Industrial Engineering, Clemson University from the Federal Aviation Administration (Program Manager: Dr. William Krebs, AAR-100). Our special thanks to Jean Watson and William Krebs from FAA for extending their support in conducting this research. We would also like to thank Rocky Ruggieri, Ken Hutcherson and the Quality Assurance department team from FedEx for their cooperation in providing data and their contribution in data gathering and interpretation sessions. The opinions, findings, conclusions and recommendations presented in this paper are those of the authors and do not necessarily reflect the views of Federal Aviation Administration.

References

1. Boeing/ATA (1995). Industry Maintenance Event Review Team. The Boeing Company, Seattle, WA.
2. FAA. (1991). Human Factors in Aviation Maintenance - Phase 1. Progress Report. DOT/FAA/AM-91/16.
3. FAA. (1993). Human Factors in Aviation Maintenance - Phase 3. Volume 1: Progress Report. DOT/FAA/AM-93/15.
4. Fitts, P.M., and Jones, R.E. (1947). Analysis of factors contributing to 460 pilot-error experiences in operating aircraft controls. Memorandum Report TSEAA-694-12, Dayton, OH: Aero Medical Laboratory, Air Material Command.
5. Gramopadhye, A.K., Drury, C.G., and Prabhu, P.V. (1997). Training for Visual Inspection. International Journal of Human Factors in Manufacturing, Volume 7(3), 171-196.
6. Hobbs, A., and Williamson, A. (2001). Aircraft Maintenance Safety Survey - Results. Department of Transport and Regional Services, Australian Transport Safety Bureau.
7. Iyengar, N., Kapoor, K., Greenstein, J.S., and Gramopadhye, A.K. (2004). Data Gathering Methodologies to Identify Impact Variables in Aviation Maintenance. Proceedings of Industrial Engineering Research Conference, May 15-19, Houston, Texas.
8. Norman, D.A. (1981). Categorization of Action Slips. Psychology Review 88, 1-15.
9. Rankin, W.L., and Allen, J. (1995). Use of the Maintenance Error Decision Aid (MEDA) to Enhance Safety and Reliability and Reduce Costs in the Commercial Aviation Industry. Proceedings of the International Air Transport Association's 1995 Aircraft Maintenance Seminar and Exhibition. The Changing Vision. November 14-16, Sydney Convention and Exhibition Center, Sydney, Australia.
10. Rankin, W., Hibit, R., Allen, J., and Sargent, R. (2000). Development and Evaluation of the Maintenance Error Decision Aid (MEDA) Process. International Journal of Industrial Ergonomics, Volume 26, 261-276.
11. Rasmussen, J. (1982). Human Errors: A Taxonomy for describing Human Malfunction in Industrial Installations. Journal of Occupational Accidents, Volume 4, 311-333.
12. Reason, J. (1990). Human Error. Cambridge University Press, New York.
13. Rouse, W.B., and Rouse, S.H. (1983). Analysis and Classification of Human Error. IEEE Transactions on Systems, Man, and Cybernetics, Volume SMC-13, No.4, 539-549.
14. Schmidt, J.K., Schmorow, D., and Hardee, M. (1998). A preliminary Analysis of Naval Aviation Maintenance Related Mishaps. Society of Automotive Engineers, Volume 107, 1305-1309.
15. Shappell, S., and Wiegman, D. (1997). A Human Error Approach to Accident Investigation: The Taxonomy of Unsafe Operations. The International Journal of Aviation Psychology, Volume 7, 269-291.
16. Shappell, S., and Wiegman, D. (2001). Applying Reason: The Human Factors Analysis and Classification System (HFACS). Human Factors and Aerospace Safety, Volume 1, 59-86.
17. Swain, A.D., and Guttman, H.E. (1983). Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications: Final Report. NUREG/CR-1278, SAND80-0200, Prepared by Sandia National Laboratories for the U.S. Nuclear Regulatory Commission.
18. Wiegman, D., and Shappell, S. (2001). A human error analysis of commercial aviation accidents using the Human Factors Analysis and Classification System (HFACS). DOT/FAA/AM-01/3, Washington DC: Federal Aviation Administration.